

ESTUDIOS PÚBLICOS

Nº 90

OTOÑO

2003

John M. Carey y Peter Siavelis

El "seguro" para los subcampeones electorales y la sobrevivencia de la Concertación

Enrique Ghersi

Barras bravas: Teoría económica y fútbol

Bárbara Eyzaguirre

Exigencias para la construcción de una prueba de selección a la universidad

Francisca Dussailant

Técnicas de medición en pruebas de admisión a las universidades

Jorge Manzi y Ernesto San Martín

La necesaria complementariedad entre teoría clásica de la medición (TCM) y teoría de respuesta al ítem (IRT)

Christina Stage

Teoría clásica de medición o teoría de respuesta al ítem: La experiencia sueca

John Finnis

Tomás de Aquino: Libertad, razón y bienes humanos

John B. Thompson

La transformación de la visibilidad

Alfredo Bryce Echenique

Del humor, del dolor, y de la risa: Crónica de una depresión

Ana María Stiven

El feminismo en retirada: A propósito del libro Right-Wing Women in Chile: Feminine Power and the Struggle against Allende 1964-1973, de Margaret Power

Óscar Godoy Arcaya

Antología del Defensor de la Paz, de Marsilio de Padua

Comisión Asesora Presidencial

Regulación del financiamiento de la actividad política

CENTRO DE ESTUDIOS PÚBLICOS

EL “SEGURO” PARA LOS SUBCAMPEONES ELECTORALES Y LA SOBREVIVENCIA DE LA CONCERTACIÓN

**John M. Carey
y Peter Siavelis**

En este trabajo se sostiene que las coaliciones de partidos en Chile, para lograr una mayoría en el Congreso, se ven en la necesidad, debido al peculiar sistema electoral de circunscripciones binominales, de inscribir a candidatos poderosos en listas precarias. Naturalmente, esto produce una divergencia entre los intereses de las coaliciones y los de los candidatos individuales. La Concertación, coalición gobernante en Chile desde 1990, habría resuelto este dile-

JOHN M. CAREY. Profesor de Ciencia Política en Washington University, en St. Louis, y profesor visitante David Rockefeller en Harvard University. Entre sus publicaciones anteriores se incluyen “Discipline, Accountability, and Legislative Voting in Latin America”, *Comparative Politics* (2003); “Parties and Coalitions in Chile in the 1990s”, en *Legislative Politics in Latin America*, editado por Scott Morgenstern y Benito Nacif (2002); “Presidential Agenda Control and Spending Policy: Lessons from General Pinochet’s Constitution” (con Lisa Baldez, *American Journal of Political Science*, 1999); “Los Efectos del Sistema Electoral sobre el Sistema de Partidos y el Respaldo Parlamentario al Ejecutivo”, *Estudios Públicos*, 55 (1994); *Presidents and Assemblies: Constitutional Design and Electoral Dynamics* (con Matthew Shugart, Cambridge University Press, 1992).

PETER SIAVELIS. Ph.D. Georgetown University. Profesor de Ciencia Política, Wake Forest University, en Winston-Salem, North Carolina. Autor, entre otras publicaciones, de *The President and Congress in Post-authoritarian Chile: Institutional Constraints to Democratic Consolidation* (Penn State University Press, 2000); “Continuity and Change in the Chilean Party System”, *Comparative Political Studies*, Vol 30, N° 6 (1997); “Nuevos Argumentos y Viejos Supuestos: Simulaciones de Sistemas Electorales Alternativos para las Elecciones Parlamentarias Chilenas”, *Estudios Públicos*, 51 (1993).

Traducción al castellano de Alberto Ide.

ma compensando con cargos en el gobierno a sus correligionarios que estuvieron dispuestos a asumir el riesgo personal, en favor de la coalición, de aceptar una candidatura y realizar una campaña vigorosa para llegar al Congreso, pero que no resultaron electos. Este sistema de “seguro” para los candidatos, sostienen los autores, le habría proporcionado a la Concertación, desde la transición a la democracia en 1990, un poderoso elemento de cohesión. Sin embargo, en vista de los estrechos resultados de la última elección presidencial (1999), la Concertación no podrá seguir garantizándoles a sus candidatos, a modo de compensación en el caso de que no salgan electos, un cargo en el aparato de gobierno. En este nuevo escenario, ¿logrará la Concertación mantener su cohesión durante el proceso de negociación de las listas de candidatos para las elecciones parlamentarias que se celebrarán en 2005?

Introducción

El fin de la Concertación se ha pronosticado repetidas veces durante los doce años en que ha gobernado. Incluso los propios dirigentes de la coalición han anunciado prematuramente su deceso (*Latin America Adviser*, 2002; *El Mercurio*, 2003a). Algunos observadores de la política chilena, por tanto, podrían considerar la disolución de la Concertación como un desenlace inevitable, producto de tantos factores, todos encaminados hacia el mismo fin, que resulta imposible asignar responsabilidades entre ellos. Compartimos la opinión de que la Concertación bien podría disolverse antes de que se celebren las próximas elecciones en Chile, pero no por las razones que comúnmente se dan. Nuestro argumento se basa, en cambio, en la estructura de las carreras políticas en Chile, las que a su vez están inextricablemente ligadas al inusual sistema electoral de distritos binominales.

Quienes presagian la extinción de la Concertación señalan que el ritmo de la economía chilena se ha vuelto algo más lento desde los primeros años de la década de 1990, cuando la coalición comenzó a gobernar. Además, el elemento que en un principio aglutinó a la coalición fue su oposición al régimen de Pinochet durante las décadas de 1970 y 1980, de manera que a medida que transcurre el tiempo la fuerza de esa motivación inicial podría debilitarse por efecto de un proceso natural. Asimismo, la Concertación, e incluso los partidos que la componen —en particular la Democracia Cristiana—, se encuentran internamente divididos en torno a problemas sociales, tales como la legalización del divorcio o el acceso al control de

la natalidad. Por último, para las elecciones de 2005 la Concertación habrá conservado la presidencia y una mayoría en la Cámara de Diputados (elegida por votación popular) por espacio de 16 años. De acuerdo con los estándares para las coaliciones multipartidistas de cualquier país —que resultan mucho menos aplicables a los sistemas presidenciales latinoamericanos—, la Concertación es una institución vetusta, marcada por las cicatrices de las acusaciones de corrupción contra algunos de sus miembros —incluido el desafuero de cinco de sus diputados en 2002— y por el desencanto generalizado que trae aparejado el hecho de llevar las riendas del poder por tanto tiempo. Así pues, podríamos concluir que la coalición gobernante se encuentra sencillamente a punto de expirar.

En verdad, cualquiera de estos factores o alguna combinación de ellos podría debilitar a la Concertación, pero ninguno de ellos nos parece necesariamente devastador para la sobrevivencia de la Concertación, por varias razones. En primer lugar, porque durante los últimos cinco años la economía chilena ha salido mucho mejor librada que la de cualquiera de sus vecinos del Cono Sur, y las cifras de empleo y crecimiento han experimentado una recuperación en los primeros meses de 2003. Hacia el año 2005 la Concertación podría perfectamente estar en condiciones de exigir un reconocimiento por su acertado manejo de la economía. Asimismo, muchos de los temas de discusión en torno al legado antidemocrático de la era Pinochet —incluidos los famosos enclaves autoritarios en la Constitución y la privilegiada situación jurídica y presupuestaria de los militares— persisten y aún son objetivos de reforma para los opositores a Pinochet. Si bien la Concertación está plagada de discrepancias internas respecto de asuntos sociales, ello ha ocurrido desde que comenzó a gobernar, y similares divisiones afectan también a sus principales oponentes en la derecha. En definitiva, ¿quién podría afirmar cuánto tiempo es demasiado para la existencia de una coalición? Si nos basamos en la experiencia histórica chilena, ya para las elecciones parlamentarias de 1997 la Concertación había excedido con creces la esperanza de vida de cualquier coalición multipartidista, y aun así tuvo éxito en esa ocasión; lo propio ocurrió en las elecciones presidenciales 1999-2000, y una vez más en los comicios parlamentarios de 2001 (Carey, 2002, pp. 222-253).

En breve, muchas de las condiciones que se consideran peligrosas para la sobrevivencia de la Concertación no son nuevas, pues éstas han estado presentes durante toda su existencia y, más bien, representan constantes del medio ambiente electoral. Además, como lo han señalado numerosos observadores, el sistema electoral chileno entraña un alto riesgo para

los partidos que abandonan las coaliciones y presentan candidatos por su cuenta (Rabkin, 1992, pp. 235-356; Siavelis, 2002, pp. 419-438). A decir verdad, la explicación que proponemos aquí respecto a la posibilidad de una desintegración de la Concertación descansa en la singularidad de la ley electoral chilena y en su interacción con los cambios recientes en las tendencias del voto. A continuación nos referiremos a estas cuestiones.

Distritos binominales y estrategias de las coaliciones

La característica más importante del sistema electoral chileno para las elecciones parlamentarias es que todas las circunscripciones eligen dos representantes¹. En consecuencia, cada lista puede incluir hasta dos candidatos en la papeleta de votación. Las listas son “abiertas”, por cuanto los votantes indican su preferencia por uno u otro candidato dentro de su lista preferida. A pesar de lo anterior, los votos obtenidos por ambos aspirantes en cada lista se acumulan primero en un total común para efectos de determinar cuántos escaños (uno, dos o ninguno) le corresponden a cada una de las listas, y luego se asignan los escaños al interior de las listas, según la votación individual de cada candidato. Los escaños se distribuyen empleando el método d’Hondt, de tal manera que la lista que alcanza el primer lugar en una circunscripción sólo puede obtener ambos escaños si logra más que el doble del total de sufragios de la lista que llegó en segundo lugar; de lo contrario, cada una de las dos listas más votadas consigue un escaño².

Pese a que el sistema electoral resulta muy claro desde el punto de vista del procedimiento, su combinación con el sistema partidista genera gran complejidad estratégica para los candidatos, los partidos y los dirigentes de las coaliciones. La política electoral en el Chile postautoritario ha estado dominada por dos grandes coaliciones —la Concertación en la izquierda y la Alianza por Chile en la derecha—, cada una de las cuales está compuesta a su vez por dos subpactos. Durante la mayor parte de este período la Concertación ha estado integrada por cuatro partidos importantes

¹ Esto se aplica a la Cámara de Diputados, cuyos 120 miembros son elegidos en 60 circunscripciones binominales, y a los miembros *elegidos* del Senado, que son votados en 19 circunscripciones binominales.

² Esta situación frecuentemente se traduce —lo que ha sido advertido en muchas descripciones de elecciones posteriores a la transición— en que el segundo candidato más votado en una circunscripción no resulte electo. Lo anterior ocurre cuando los candidatos que alcanzan el primer y segundo lugar pertenecen a la misma lista, pero el total de votos de dicha lista no dobla el total de sufragios de la lista que ha logrado el segundo lugar.

(PDC, PRSD, PPD, PS), y la Alianza por Chile por tres (RN, UDI, UCC)³. Los independientes vinculados con determinados sectores ideológicos desempeñan una importante función y también han tenido que negociar su inclusión en las principales listas electorales.

Los partidos que forman cada coalición deben negociar 60 listas de candidatos para la Cámara de Diputados, y 9 ó 10 listas de candidatos para el Senado, lo que depende del ciclo de rotación en la cámara alta. El patrón observado en elecciones anteriores indica que hasta ahora cada una de las coaliciones ha asignado un escaño, en cada circunscripción, a cada uno de sus dos subpactos, junto con hacer ajustes en la conformación de las listas, aquí y allá, con el fin de atraer a pequeños partidos que pueden ayudar a reforzar la votación de la coalición a nivel nacional en las elecciones presidenciales.

Dos crudas realidades dan forma a las negociaciones. En primer lugar, la mayoría de las listas están expuestas a sufrir una derrota en cada circunscripción porque el umbral para obtener los dos escaños es demasiado alto. La Tabla 1 muestra el número de veces que cada una de las principales coaliciones logró “doblar” (es decir, ganar los dos escaños de una circunscripción) en elecciones recientes. En segundo lugar, y muy importante para nuestros fines, la formación de la dupla de candidatos que irá en cada lista, en cada circunscripción, es crucial. La clave de la victoria para los subpactos y partidos individuales radica en que el compañero de lista del candidato de su colectividad sea una figura sumamente fuerte, que le permita a la coalición conseguir más del doble del total de votos de la lista que llegue en segundo lugar, o bien que el compañero sea una figura relativamente débil que no obtenga más sufragios que el candidato del propio partido, pero donde la coalición consiga suficientes votos para lograr un escaño. Este delicado equilibrio crea una tensión entre lo que les interesa a los candidatos y lo que les interesa a las coaliciones. Las coaliciones quisieran formar 60 duplas destinadas a maximizar los votos, en tanto que los candidatos por lo general preferirán un compañero de lista débil⁴. A los candidatos que son superados en cantidad de sufragios por su compañero de lista los denominamos “subcampeones”. Más adelante analizaremos cuál es la suerte que corren.

³ La Alianza por Chile también ha sido conocida como “Democracia y Progreso” y “Unión por el Progreso”. La UCC se alineó formalmente con la derecha en la elección de 1993.

Explicación de las siglas de los partidos:

PDC: Partido Demócrata Cristiano; PRSD: Partido Radical Social Demócrata; PPD: Partido Por la Democracia; PS: Partido Socialista; RN: Renovación Nacional; UDI: Unión Demócrata Independiente; UCC: Unión de Centro-Centro.

⁴ Para un análisis más exhaustivo del intrincado funcionamiento del sistema partidista chileno y de las condiciones empíricas que afectan a la formación de coaliciones, véase Siavelis (2002).

TABLA 1: RESULTADOS DE COMPETENCIAS ENTRE COALICIONES A NIVEL DISTRITAL EN LAS ELECCIONES PARLAMENTARIAS CHILENAS, 1989-1997

	Derecha dobla	Concertación			Subcampeones de la Concertación	
		Dobla	No gana ninguno de los escaños	Gana un escaño (el otro escaño lo obtuvo un candidato que no es de la derecha)		
Cámara de Diputados	1989	0	11	2	1	46
	1993	1	11	1	0	48
	1997	0	9	0	4	47
Senado	1989*	0	3	0	0	16
	1993	0	0	0	0	9
	1997	0	1	0	0	9

* En 1989 orrespondió elegir a los 38 senadores sujetos a elección popular, en un total de 19 circunscripciones. En las elecciones posteriores se eligieron cada vez 19 senadores en 9 ó 10 circunscripciones, ya que el Senado se renueva parcialmente cada 4 años.

Coaliciones, mayorías y el riesgo del “doblaje”

El sistema chileno de circunscripciones binominales plantea el complejo problema de distribuir el riesgo en circunstancias en que los objetivos colectivos de las coaliciones electorales entran en conflicto con las carreras políticas de los candidatos individuales. Consideremos la siguiente situación: las coaliciones desean maximizar la cantidad de escaños que controlan en el Congreso, y en particular quisieran ser mayoría parlamentaria. Una coalición no puede tener mayoría en el Congreso si no consigue “doblar” por lo menos en una circunscripción. Teniendo en cuenta el deseo de obtener la mayor cantidad posible de escaños, una coalición ambiciosa se ve en la necesidad de procurar doblar a las listas contendoras en tantas circunscripciones como lo permita su fuerza electoral. Así pues, las coaliciones que aspiran a ser mayoría en el Congreso se esfuerzan en identificar circunscripciones donde sea posible doblar, y tratan de orientar hacia ellas recursos suficientes como para cruzar el umbral crítico.

Uno de los recursos más importantes que las coaliciones pueden distribuir entre las circunscripciones son los propios candidatos. Durante la primera década después de la transición a la democracia, una lista de la Concertación cuyo desempeño resultaba particularmente contundente podía abrigar la razonable esperanza de doblar a la principal lista opositora, mien-

tras que esta última sólo tenía probabilidades de doblar los votos de su adversario en circunstancias extremadamente propicias⁵. El “doblaje” ha tenido, por tanto, un carácter estratégico principalmente para la Concertación, cuyas probabilidades de seguir logrando ese objetivo dependen de contar con candidatos que puedan exhibir un desempeño sobresaliente, proporcionando el impulso necesario para que la lista pueda cruzar el umbral mágico. El problema es que, en estas circunstancias, la meta colectiva de la coalición entra en directo conflicto con el objetivo político más inmediato de la mayoría de los candidatos a un cargo electivo: la sobrevivencia política.

Los candidatos quieren salir electos, ya sea por mera ambición personal, ya sea por el deseo más noble de formular políticas públicas adecuadas, o bien por una combinación de ambos. Si pierden, su capacidad de cumplir cualquiera de esos objetivos quedará menoscabada. Además, son precisamente los candidatos más poderosos —aquellos que por sus cualidades personales y popularidad son los más aptos para ganar las elecciones— a quienes las coaliciones deben exponer a las situaciones más riesgosas para aspirar al objetivo colectivo de doblar la votación de la lista adversaria, y para obtener mayoría en el Congreso.

Un candidato poderoso tiene razones de sobra para preferir un compañero de lista relativamente débil cuando se trata de una circunscripción donde su coalición está dispuesta a conformarse con un escaño. Ser el candidato de peso en una lista de ese tipo constituye, en efecto, garantía de salir elegido. Por otra parte, ser candidato en una circunscripción donde la coalición a la que se pertenece aspira a doblar, significa no sólo que se deberá enfrentar la competencia de otras coaliciones sino además la de su propio compañero de lista. El mismo imperativo que mueve a las coaliciones a presentar duplas de candidatos poderosos en circunscripciones donde se pretende doblar es el que amenaza la seguridad electoral de sus mejores políticos.

Seguro contra el riesgo: Asignación de cargos en el gobierno

¿De qué manera ha resuelto la Concertación este dilema? Suponemos que con el fin de inducir a los candidatos fuertes a asumir el riesgo de una campaña para doblar, la coalición gobernante ha ofrecido un seguro.

⁵ Los porcentajes promedio de los votos obtenidos por ambas coaliciones fueron, respectivamente, el 59% y el 41%, con una desviación estándar del 9%. De manera que la Concertación podía doblar la votación de su adversario con un desempeño de una desviación estándar superior al promedio, mientras que el rendimiento de la derecha tenía que superar el promedio en tres desviaciones estándar.

Para usar la terminología de la economía política, por más de una década la Concertación ha cosechado sustanciales ganancias gracias a su capacidad de actuar colectivamente, y esas “ganancias” no son otras que el control del gobierno: tanto de los cargos más altos en el Congreso, debido a que la Concertación ha logrado doblar la votación de sus oponentes en circunscripciones clave, como del Poder Ejecutivo, debido a su capacidad para nominar candidatos presidenciales y apoyarlos de manera cohesionada. Para los candidatos parlamentarios fuertes que corren riesgos en beneficio de la coalición al tratar de doblar en una circunscripción, el seguro adquiere la forma de la promesa de ser nombrados en atractivos cargos en el gobierno si alguno de ellos no alcanzara el objetivo previsto en la justa electoral. En términos muy simples, el trato es el siguiente: “Usted corre un riesgo en esta elección por el bien de la coalición. Si ganamos ambos escaños, magnífico. Si no conseguimos doblar los votos de la lista contendora y usted es derrotado, se le compensará con un ministerio o la asignación de otro cargo atractivo en el Poder Ejecutivo”.

Modelo

Proporcionamos a continuación un modelo de las condiciones en que a los candidatos que alcanzan el segundo lugar en las listas de la Concertación que no logran doblar, se les asigna tras su derrota electoral un cargo en el gobierno. A estos candidatos que no lograron salir electos los llamamos “subcampeones”. El modelo es el siguiente:

$$\begin{aligned} \text{Pr (Asignación de Cargo)} = & a \text{ (Constante)} \\ & + b1 \text{ (Cámara de Diputados)} \\ & + b2 \text{ (Razón Coalición)} \\ & + b3 \text{ (Razón Concertación)} \end{aligned}$$

donde:

- *Asignación de Cargo* adopta el valor de 1 si el subcampeón fue designado ministro, subsecretario o embajador durante el período legislativo posterior a su fallida candidatura al Congreso.
- *Cámara* adopta el valor de 1 para los candidatos a la Cámara de Diputados y de 2 para los candidatos al Senado.
- *Razón Coalición* es el total de votos para la lista de la Concertación dividido por el total de votos para la lista de la Alianza por Chile.

- *Razón Concertación* corresponde a los votos para el candidato que alcanza el primer lugar al interior de la lista de la Concertación divididos por los votos conseguidos por el candidato que ocupa el segundo lugar en dicha lista.

Las candidaturas senatoriales son más importantes que las candidaturas a la Cámara de Diputados, porque si bien ambas cámaras cuentan con poderes análogos, existen menos escaños en la cámara alta, con lo cual cada uno de éstos se torna máspreciado. El Senado también es percibido como la plataforma de lanzamiento de las candidaturas presidenciales. Puesto que hay más en juego en las elecciones senatoriales, la lucha es más reñida y un desempeño sobresaliente resulta más impresionante. Asimismo, los candidatos reclutados para competir en la carrera hacia el Senado son individuos más prominentes, lo cual tiene dos implicaciones. En primer lugar, su experiencia y su preparación indican que por lo general tienen “calibre ministerial”. En segundo lugar, los candidatos al Senado que poseen altas credenciales cuentan con atractivas opciones fuera del ámbito político, de modo que para inducirlos a emprender una campaña riesgosa el “seguro” para los subcampeones tiene que ser más generoso que el ofrecido a los candidatos a diputados. Por todas estas razones esperamos que el coeficiente en la Cámara de Diputados sea positivo.

La Razón Coalición es más alta cuanto más cerca se encuentra la Concertación de doblar la votación de la derecha⁶, de manera que cuanto mayor es la Razón Coalición, más impresionante será el desempeño colectivo de la lista en su intento por lograr el doblaje. Lo que a nuestro juicio cabría esperar es que los subcampeones provenientes de esas listas sean recompensados por su aporte al esfuerzo colectivo, de suerte que el coeficiente en la Razón Coalición debería ser positivo.

La Razón Concertación es más baja cuanto más equiparados se encuentren los dos candidatos de la lista de la Concertación. Si el excelente desempeño de una lista es en gran medida producto de la capacidad de un solo candidato para atraer votos, entonces el subcampeón no sería considerado como un instrumento que contribuye al desempeño general de la lista, y por ende tampoco como merecedor de una recompensa postelectoral, como cuando ambos candidatos de la lista corren nariz a nariz. Siguiendo esta lógica, el coeficiente en la Razón Concertación debería ser negativo. Sin embargo, esta hipótesis nos inspira menos confianza que las variables

⁶ Las listas de la Concertación que logran doblar en una circunscripción ciertamente no producen subcampeones, ya que sus dos candidatos resultan elegidos. Por ello, esas listas no se incluyen en el análisis.

Cámara de Diputados y Razón Coalición, porque el hecho de fomentar la competencia al interior de la lista no necesariamente redundará en beneficio para la coalición. Es decir, las coaliciones se benefician cuando tienen candidatos de peso que, en un sentido genérico, atraen votos para la lista, pero no se benefician de la competencia mediante la cual los compañeros de lista procuran quitarse votos entre sí. En efecto, una competencia demasiado intensa y mutuamente destructiva podría resultar perniciosa para el posicionamiento general de la lista y hacer bajar su votación general, del mismo modo que a los partidos estadounidenses les preocupan las campañas primarias agresivas, ya que éstas a la larga perjudican al candidato que sale finalmente nominado. Entonces, con el fin de reducir los incentivos para una competencia “canibal” al interior de la lista, la coalición tal vez desee hacer hincapié en el desempeño colectivo de la lista ofreciendo un “seguro” para los que salgan subcampeones.

En suma, estimamos que las probabilidades de que a los subcampeones de la Concertación se les asigne un cargo serán mayores:

- cuando se trata de candidaturas para el Senado y no para la Cámara de Diputados;
- cuando la lista de la Concertación inflija a la lista contendora una derrota más categórica (aproximándose al “doblaje”); y
- (con menos confianza en este caso) cuando los dos candidatos de la Concertación contribuyan más equitativamente a la votación total de la lista.

Nosotros interpretaríamos esos resultados como evidencia de que la política de “seguros” mediante cargos asignados por el Ejecutivo se aplicó sistemáticamente como una manera de atraer a candidatos poderosos para que participaran en campañas reñidas, y de recompensar a quienes actuaron con coraje defendiendo una causa perdida.

A estas alturas vale la pena subrayar lo que son, a nuestro juicio, tanto los aspectos relevantes como las limitaciones de nuestro análisis estadístico. Nuestro modelo mide si, entre los subcampeones, los candidatos más poderosos tienen más posibilidades que los candidatos más débiles de recibir un nombramiento. Nuestra medición de la fortaleza de un candidato se basa enteramente en los resultados electorales, los cuales reflejan combinaciones de dos tipos de características:

1. calidad inherente (carisma, calificaciones, relieve personal), y
2. esfuerzo en beneficio de la campaña.

La lógica por la cual estas dos cualidades contribuyen a la probabilidad de recibir un cargo es diferente. Nosotros esperamos que la fortaleza

inherente de un candidato potencial elevará los costos de oportunidad asociados al hecho de aceptar una candidatura riesgosa, porque las mismas cualidades que hacen que un candidato sea un competidor inherentemente poderoso aumentan, asimismo, las probabilidades de que tenga a su disposición opciones de carrera más atractivas que la de exponerse a un riesgo electoral. Para tentar a este tipo de candidatos a aceptar que sean incluidos en una lista de alto riesgo (que procura doblar a la lista adversaria) resulta fundamental poder prometerles que contarán con un seguro. Con respecto al esfuerzo desplegado por un candidato, la lógica es más directa. Asimismo, el Poder Ejecutivo debería demostrar una mayor disposición a recompensar a aquellos que trabajaron arduamente en beneficio de la coalición y que les faltó muy poco para doblar en su circunscripción, que a aquellos que realizaron una campaña menos eficaz y en consecuencia estuvieron mucho más lejos de lograr ese objetivo. Si bien en este caso operan dos lógicas distintas, ambas generan las mismas expectativas con respecto a las estimaciones de los parámetros en nuestro modelo.

Lo ideal sería medir por separado la calidad inherente de un candidato y el esfuerzo que éste ha realizado, para así estimar los efectos independientes de cada característica. Si pudiéramos hacerlo, entonces lograríamos establecer una comparación pertinente entre candidatos electos de la Concertación que compitieron en circunscripciones en que la lista sólo obtuvo uno de los dos escaños (es decir, aquellas donde no se dobló) y aquellos candidatos que resultaron subcampeones en distritos donde el objetivo de la lista de la Concertación era doblar la votación de la lista adversaria. Si estamos en lo correcto, entonces la calidad inherente de los subcampeones que hicieron campaña en distritos donde la meta era doblar, debería ser aun mayor que la de los ganadores en toda la serie de distritos donde sólo fue elegido un candidato de la Concertación. Desgraciadamente, aparte de los resultados electorales no tenemos otra manera de medir la calidad inherente de los candidatos. Por consiguiente, debemos reconocer que nuestro modelo no nos permite separar la “calidad inherente/costos de oportunidad” respecto al seguro, de la “recompensa por el esfuerzo”. A decir verdad, ambos aspectos son sin duda mutuamente coherentes, y a nuestro juicio ambos tienen algo de verdad.

Los datos

Los datos en que se basan las variables independientes son de acceso público y pueden recopilarse fácilmente. En el sitio web del Ministerio del Interior están disponibles los resultados electorales, a nivel de circuns-

cripciones, de todas las elecciones parlamentarias desde 1989 —en cifras desglosadas por candidato, partido y lista de coalición⁷—.

Por otra parte, la recopilación de datos sobre cargos asignados supone una labor más intensa, y además estos antecedentes son más difíciles de medir de manera integral. Aplicamos dos estrategias para reunir los datos de las asignaciones de cargos. Primero creamos una variable, Asignación de Cargo, siguiendo tres etapas:

1. Identificamos a los subcampeones de la Concertación, que son los candidatos de la coalición derrotados, *excepto* cuando se dan las siguientes condiciones:

- por supuesto, cuando la Concertación dobló la votación de la derecha, eligiendo así a ambos candidatos de su lista (es decir, no hubo perdedores en la lista de la Concertación);
- la lista de la Concertación incluyó un solo candidato, quien obtuvo un escaño;
- la Concertación no ganó ninguno de los dos escaños, sufriendo así una humillación electoral en la circunscripción; o
- la Concertación ganó el primer escaño, pero el segundo fue obtenido por el candidato de una lista que no era de la derecha, de manera que la Concertación ni siquiera logró doblar compitiendo con una oposición dividida, circunstancia que debe ser la más propicia para conseguir ese objetivo. (La Tabla 1, más arriba, muestra la frecuencia de cada uno de estos escenarios, al igual que la restante cantidad de campañas en cada elección, las cuales generaron subcampeones, de modo que forman parte de nuestro conjunto de datos).

2. Recopilamos los nombres de personas designadas en una serie de cargos políticos entre 1990 y 2001 (lapso que abarca los períodos parlamentarios posteriores a las tres elecciones que analizamos).

3. Cotejamos los nombres de nuestros subcampeones con los que ocupaban los cargos que habíamos identificado, y asignamos a la variable Asignación de Cargo el puntaje de 1 si el subcampeón recibía un nombramiento compensatorio en un alto cargo durante el período parlamentario posterior a su derrota electoral, y 0 si ello no ocurría.

En la Tabla 2 se listan los cargos que revisamos para construir la variable Asignación de Cargo, y nuestras fuentes. En Chile, el Primer Mandatario cuenta con amplios poderes de designación, ya que nombra a

⁷ <http://www.elecciones.gov.cl/full/indexf.htm>

TABLA 2: CARGOS DISPONIBLES INCLUIDOS EN LA VARIABLE ASIGNACIÓN DE CARGO (Datos para el período 1990-2001)

Cargo	Número de cargos disponibles	Fuentes de los datos
Ministro	21	Keesings
Subsecretario	30	Ministerio Secretaría General de la Presidencia, División Coordinación Interministerial
Embajador	82 ^a	Ministerio de Relaciones Exteriores

^a Número de embajadores designados por el Presidente entre 1990 y 2001, de acuerdo con datos proporcionados por el Ministerio de Relaciones Exteriores.

aproximadamente 3.500 funcionarios gubernamentales (*El Mercurio* 2003b). Sin embargo, nos concentramos sólo en los cargos de alto perfil. Como los puestos de menor nivel resultan menos atractivos para los políticos ambiciosos, disminuyen las posibilidades de usarlos como recompensa. Por cierto que no incluimos la totalidad de los nombramientos importantes. Hay otros cargos con los que los jefes de Estado podrían recompensar razonablemente a los políticos que han luchado con coraje en las batallas electorales libradas por su coalición, pero de los cuales no logramos obtener datos sistemáticos (por ejemplo, miembros de directorios de empresas estatales, gobernadores regionales). Además de los anteriores, hay otros nombramientos (por ejemplo en el Poder Judicial, en los partidos, futuras nominaciones en listas de candidatos para el Congreso o para alcaldías) que no son de exclusivo resorte del Poder Ejecutivo pero en los que el Primer Mandatario puede ejercer una considerable influencia. En resumen, la variable Asignación de Cargo incluye casi con certeza algún error del Tipo II, pero no del Tipo I. Vale decir, se puede afirmar casi con seguridad que no logramos identificar a algunos subcampeones que fueron compensados con cargos por el Presidente de la República, o a instancias de éste, pero estamos seguros de que todos aquellos que identificamos como beneficiarios de nombramientos están correctamente clasificados.

Nuestra segunda estrategia para recopilar datos sobre Asignación de Cargo complementa la primera al minimizar el error del Tipo II (no detectar nombramientos que *fueron hechos*) mediante exhaustivos esfuerzos para trazar la trayectoria seguida por los subcampeones después de su derrota electoral. Como no podíamos investigar con tanta acuciosidad a todos los candidatos incluidos en el estudio, limitamos esta estrategia a los subcampeones de la Concertación que, de acuerdo con nuestras hipótesis, eran los aspirantes con *más y menos* probabilidades de ser compensados con nom-

bramientos. Identificamos como “aspirantes más probables” a los subcampeones de las listas en que la razón de votación Concertación:Derecha fue mayor que 1,8 —esto es, la lista obtuvo el 90% de los votos necesarios para lograr doblar a la derecha, faltándole muy poco para cumplir su objetivo—. Se detectaron 20 de esos subcampeones: 14 candidatos a la Cámara de Diputados y 6 al Senado. Identificamos a los “aspirantes menos probables” como los subcampeones de las listas de la Concertación que recibieron *menos* votos que la derecha (razón Concertación:Derecha entre 0,5 y 1,0). Ello arrojó un total de 19 subcampeones para la Cámara de Diputados y sólo uno para el Senado. Con el fin de no excluir a los candidatos al Senado de la lista de los aspirantes menos probables, añadimos a los siguientes cuatro subcampeones de la Concertación que postularon a la Cámara alta (con razones de hasta 1,13)⁸.

Para nuestro conjunto de “aspirantes más y menos probables” complementamos los antecedentes incluidos en la variable Asignación de Cargo con búsquedas estándar en internet, búsquedas manuales en sitios web y archivos del gobierno chileno y de partidos, diarios y ONG de este país, y correspondencia por correo electrónico con académicos y funcionarios gubernamentales de Chile para indagar acerca de las carreras postelectorales de determinados subcampeones. Este enfoque suministró información sobre algunos nombramientos del Poder Ejecutivo para los cuales no disponemos de datos exhaustivos (es decir, aquellos no incluidos en la variable Asignación de Cargo), y sobre nombramientos en otros puestos en que el apoyo presidencial puede haber resultado determinante⁹.

⁸ La incorporación de estos subcampeones que postularon al Senado debería predisponer en contra de nuestro modelo, ya que estamos fortaleciendo la categoría de “aspirantes menos probables” con candidatos al Senado, los cuales tienen mejores perspectivas que los postulantes a la Cámara de Diputados para conseguir nombramientos.

⁹ Los aspirantes “más probables” y “menos probables” fueron objeto de una exhaustiva indagación, que comenzó con nuestra base de datos de nombramientos, seguida de una amplia búsqueda en todo el sitio web “Gobierno de Chile” (incluidos los parlamentarios en ejercicio y los anteriores). Proseguimos con búsquedas combinadas de nombres de pila, apellidos y partido político, luego nombres de pila y apellidos, y luego simplemente apellidos, utilizando el buscador Google (reconocido ampliamente como el más completo y preciso). Examinamos cada coincidencia e hicimos un seguimiento para asegurarnos de haber identificado a la persona correcta, cotejándola con el correspondiente puesto ocupado después de las elecciones. Se podría argumentar que la eventual mayor visibilidad de los “aspirantes más probables” aumenta las posibilidades de encontrarlos, lo cual inclina los resultados en favor de nuestra hipótesis. En la medida en que esto ocurrió, nuestros esfuerzos de búsqueda fueron sin embargo asimétricos, con una mayor cuota de dedicación a los “aspirantes menos probables”, pues la información sobre su carrera postelectoral era generalmente más escasa. En fin de cuentas, confiamos en que la información que recopilamos con precisión refleja la suerte corrida por ambos grupos después de las elecciones.

Resultados

Como primera aproximación al tema de si los subcampeones de la Concertación que exhibieron un buen desempeño en beneficio de la coalición fueron recompensados sistemáticamente con nombramientos en cargos, consideremos los resultados de las indagaciones exhaustivas sobre nombramientos a “los aspirantes más y menos probables”, que se resumen en la Tabla 3. Identificamos nombramientos en cargos apetecibles para el 70% de nuestros subcampeones en la categoría de “aspirantes más probables”, pero sólo para el 25% de nuestros subcampeones en la categoría de “aspirantes menos probables”. Entre ellos se incluyen nombramientos realizados por el Presidente durante el período inmediatamente posterior a la elección respectiva, los cuales beneficiaron a 3 de nuestros 6 “aspirantes más probables” que postularon al Senado y terminaron como subcampeones, además de puestos importantes en la administración pública, no dependientes directamente del Primer Mandatario, que favorecieron a otros dos más, y una nominación para una (exitosa) campaña senatorial con que se recompensó al sexto. Así pues, la totalidad de los 6 subcampeones en elecciones senatoriales con más probabilidades de lograr nombramientos obtuvieron alguna compensación. Entre los cinco subcampeones en elecciones senatoriales con menos probabilidades de obtener nombramientos identificamos a uno de ellos en un cargo asignado por el Presidente durante el período posterior a la elección y a otro durante el segundo período posterior a la elección, y ningún cargo para los otros tres. Existen más subcampeones en las elecciones para la Cámara de Diputados, pero el patrón es el mismo: identificamos cargos asignados durante el período posterior a la elección que benefició a la mayoría de los aspirantes más probables, pero a ninguno de los aspirantes menos probables.

Por cierto, cabría esperar que el método empleado para escoger las carreras políticas que serían indagadas en profundidad inevitablemente conduciría a esos resultados tan crudos. Lo que podemos concluir en esta etapa es que aquellos subcampeones que tenían, según hemos sostenido, las perspectivas más auspiciosas después de sufrir una derrota electoral, cuentan con buenas posibilidades de recibir nombramientos compensatorios en altos cargos (incluso tal vez mayores probabilidades que las sugeridas en la Tabla 3 si se tiene en cuenta que se nos pueden haber escapado algunos cargos asignados), mientras que para aquellos subcampeones cuyas expectativas nos parecieron menos promisorias, las perspectivas eran bastante más sombrías.

Habiendo salvado este primer umbral sometimos la teoría a una prueba estadística más amplia utilizando los datos electorales a nivel de

TABLA 3: CARRERAS POLÍTICAS POSTELECCIONES, EN BASE A INDAGACIONES EXHAUSTIVAS DE LOS SUBCAMPEONES CONSIDERADOS COMO ASPIRANTES MÁS Y MENOS PROBABLES A RECIBIR NOMBRAMIENTOS, SEGÚN LAS RAZONES DE VOTACIÓN DE SUS LISTAS CONCERTACIÓN: DERECHA (Los subcampeones en elecciones senatoriales aparecen destacados en *cursivas*)

Nombramiento posterior a la elección	Aspirantes más probables	Aspirantes menos probables
Cargos asignados por el Presidente durante el período posterior a la derrota de los subcampeones	<i>Jara Wolff</i> <i>Lagos Escobar</i> <i>Rebolledo González</i> Saavedra Cortés Aylwin Oyarzún Pérez Lobos Molina Valdivieso	<i>Estévez Valencia</i> Auth Stewart
Cargos asignados por el Presidente durante el período subsiguiente al de la derrota de las subcampeones	Yunge Bustamante Santelices Altamirano	<i>Correa Díaz</i>
Nominación para candidatura con miras al próximo período parlamentario	<i>Zaldívar Larraín</i>	Girardi Lavín Romero Fuentes
Cargo político no asignado por el Presidente	<i>Sule Candia^a</i> <i>Navarrete Betanzo^b</i> Arenas Escudero ^c Dintrans Schafer ^d	Torres Gutierrez ^e
Desconocido	Gajardo Chacón Morales Abarzúa Del Valle De La Cruz Sepúlveda Gutiérrez Vargas Vega Leblanc Valenzuela	<i>Vodanovic Schnake</i> <i>Sáenz Rojas</i> <i>Carmine Zúñiga</i> Campos Leal Esquivel Santander Álvarez Pinto García Villarroel Arancibia Silva Gustavo Núñez Witker Velásquez Fuentes Araya Calquín Monardes Alarcón Rojas Rossetti Gallardo Molina Sepúlveda Ibáñez Vergara Apara Álamo Makluf Campos

^a Presidente del Partido Radical después de la elección de los subcampeones; fallecido.

^b Presidente del Comité Olímpico de Chile.

^c Secretario General del PPD a contar de 2000.

^d Abogado integrante de la Corte de Apelaciones de Rancagua: cargo asignado por el Poder Judicial y no por el Poder Ejecutivo.

^e Nominado por el partido para presentarse como candidato a concejal en 2002 (fue elegido).

circunscripción y la variable Asignación de Cargo, basándonos únicamente en nuestras búsquedas exhaustivas de aquellos cargos enumerados en la Tabla 2. Como la variable dependiente es binaria (0:1), el modelo es una regresión logística. Los resultados básicos, junto con algunos valores que facilitan la interpretación, se encuentran en la Tabla 4.

La probabilidad de conseguir un nombramiento depende en gran medida de la variable Cámara de Diputados y del desempeño de la lista, aunque no de la paridad relativa de los candidatos al interior de una determinada lista de la Concertación. Los coeficientes para las tres variables independientes adoptan los signos esperados: positivo para Cámara de Diputados y la Razón Coalición, y negativo para la Razón Concertación. Los dos primeros son estadísticamente significativos en los niveles convencionales, mientras que el último no alcanza el nivel esperado, lo cual sugiere que determinados subcampeones al interior de las listas de la Concertación son recompensados de acuerdo con el desempeño colectivo de la dupla de candidatos y no según el total de votos alcanzados individualmente.

Las últimas dos columnas de la Tabla 4 muestran el efecto marginal estimado del cambio en Cámara de Diputados y en la Razón Coalición, en cada caso con otras variables fijadas en sus niveles medios, de acuerdo con simulaciones efectuadas empleando el software Clarify (King, Tomz, y

TABLA 4: RESULTADOS DE LA REGRESIÓN LOGÍSTICA DE LA VARIABLE DEPENDIENTE ASIGNACIÓN DE CARGO EN CÁMARA DE DIPUTADOS, RAZÓN COALICIÓN Y RAZÓN CONCERTACIÓN

Variable dependiente: Asignación de Cargo				
Variables independientes	Coeficiente (error estándar)	P > z	Aumento de probabilidad de recibir un nombramiento	Intervalo de confianza de 95%
Constante	-5,99 (1,72)	0,00		
Cámara de Diputados	1,33 (0,53)	0,01	Senado, en relación con Cámara de Diputados: +15%	+3% — +32%
Razón Coalición	1,60 (0,92)	0,08	Cuando la Razón Coalición entre subcampeones va desde su percentil 20 (1,1) hasta su percentil 80 (1,7): +8%	0% — +17%
Razón Concertación	-0,11 (0,20)	0,59		
N=167				
Pseudo R ² =0,10				

Wittenberg, 2000, pp. 341-355). Asimismo, el modelo estima que las probabilidades de ser nombrado en un cargo por el Presidente después de las elecciones son 15% mayores para los subcampeones que postulan al Senado que para los que pretenden llegar a la Cámara de Diputados, y si bien la estimación específica es incierta, la última columna muestra que podemos estar confiados en que el efecto es positivo y considerable. Del mismo modo, manteniendo la variable Cámara de Diputados constante, las perspectivas de obtener un nombramiento son 8% mayores para los subcampeones de listas que tuvieron un buen desempeño (percentil 80, de acuerdo con la razón de votación Concertación:Alianza por Chile) que para aquellos incluidos en listas que obtuvieron una baja votación (percentil 20), y una vez más podemos estar ciertos de que el efecto es positivo. Dado que la probabilidad general de conseguir un nombramiento es de sólo el 9% entre todos nuestros subcampeones (15/167), la diferencia entre formar parte de una lista de la Concertación que casi logra doblar e integrar una lista que compite de manera muy estrecha con la derecha, duplica, aproximadamente, las posibilidades del subcampeón de lograr un nombramiento.

Al interpretar estas cifras es importante tener en cuenta que los datos faltantes en la variable Asignación de Cargo deberían debilitar nuestras estimaciones de parámetros. Suponemos que se han quedado cortas nuestras estimaciones de los efectos marginales de Cámara de Diputados y de la Razón Coalición en la probabilidad de conseguir un nombramiento.

El sistema electoral chileno genera una tensión entre los incentivos para las coaliciones y los alicientes para los candidatos individuales. Con el objeto de alinear los incentivos para los políticos con aquellos destinados a la coalición en su conjunto, la Concertación ha proporcionado un “seguro” a sus subcampeones. La conclusión que se desprende de nuestro análisis es que aquellos que fueron reclutados para participar en las campañas más prominentes (es decir, para llegar al Senado) y aquellos cuyos esfuerzos estuvieron más cerca de producir frutos (es decir, cuyas listas casi lograron doblar la votación de sus oponentes de la derecha, contribuyendo así a que la Concertación alcanzara una mayoría parlamentaria) fueron los principales beneficiarios de las compensaciones con cargos. La evidencia sugiere, entonces, que la Concertación ha utilizado los nombramientos en cargos políticos para conciliar los intereses individuales de los políticos con los de la coalición en su conjunto —recompensando a los políticos que aceptaron correr un riesgo personal en beneficio de la coalición y que contribuyeron a mejorar el desempeño colectivo de sus listas—, y superar así la divergencia de incentivos generada por el peculiar sistema electoral chileno.

El nuevo panorama electoral chileno

El sistema de compensaciones ha funcionado adecuadamente para la Concertación y para sus candidatos en cada una de las cuatro elecciones celebradas desde la transición a la democracia en Chile. ¿Por qué debería fracasar ahora? Sugerimos que el “fondo de compensación” ya no es seguro, puesto que éste requiere el control del Poder Ejecutivo. Desde la perspectiva de los candidatos que podrían conseguir doblar los votos del adversario, el peligro es que bien puede ocurrir que la Concertación no controle el Poder Ejecutivo tras la elección de 2005. Esta posibilidad se tornó evidente en los comicios celebrados en 1999 y 2000, por el hecho de que en la lucha por la Presidencia de la República Joaquín Lavín resultó para Ricardo Lagos un contendor mucho más poderoso que lo previsto. Desde esa elección Lavín ha trabajado arduamente para posicionarse como el candidato que volverá a presentarse en las elecciones de 2005, y ya en este momento encabeza claramente los sondeos. De acuerdo con una encuesta realizada por el Centro de Estudios Públicos en diciembre de 2002, el 40% de los entrevistados mencionó a Lavín cuando se les formulaba la siguiente pregunta abierta: “¿Quién desearía que fuera el próximo Presidente de la República?” Quien le sigue en la lista de candidatos favoritos, la ministra de Relaciones Exteriores del actual gobierno de la Concertación, Soledad Alvear, sólo fue nombrada por el 9% de los encuestados. El 56% opinaba que Lavín sería el próximo Presidente y el 53% manifestó su voluntad de votar por él (Centro de Estudios Públicos, 2003).

Al momento de redactar este trabajo faltan casi tres años para la elección de 2005, y las perspectivas de que la Concertación conserve en sus manos la presidencia bien pueden mejorar durante ese lapso. Con todo, la situación es fundamentalmente distinta de la que ha precedido todas las elecciones parlamentarias desde la transición hasta hoy. En cada uno de los casos anteriores, no había dudas de que la Concertación mantendría el control del Poder Ejecutivo y de todos los cargos que dependen de esta rama del gobierno.

- En 1989, la victoria de Patricio Aylwin era prácticamente incuestionable, tras encabezar las fuerzas unificadas del “No” en el plebiscito de 1988 y liderar la oposición al gobierno de Pinochet durante la transición posterior.
- En 1993, luego de que se llegó a un acuerdo al interior de la Concertación en torno a la candidatura de Eduardo Frei (*antes* de que se elaborasen las listas de candidatos para las campañas parlamentarias), nunca hubo duda acerca de cuál sería el resultado de la elección presidencial.

- Más aun, el establecimiento de un período presidencial de 6 años significó que las elecciones parlamentarias de 1997 se celebraran cuando habían transcurrido dos años del gobierno de Frei, tiempo suficiente para proporcionar considerables compensaciones a los subcampeones merecedores de gratificaciones en la forma de nombramientos.
- Por último, la siguiente serie de elecciones parlamentarias tuvo lugar en 2001, dos años después de que la figura de Lavín se revelara como una manifiesta amenaza para la presidencia, pero cuando al Presidente Lagos le restaban aún cuatro años de permanencia en el poder, tiempo suficiente para compensar con nombramientos a los subcampeones de la Concertación.

En 2005 la situación será distinta. A menos que la Concertación logre restablecer el predominio del que disfrutó a fines de la década de 1980 y principios de los años noventa en cuanto a respaldo electoral, la capacidad de la coalición para pagarles un “seguro” a los candidatos que no logren doblar estará en tela de juicio. Es más, en 2005 las elecciones parlamentarias y presidenciales coincidirán por primera vez desde 1993. Las elecciones legislativas de 1997 y 2001 se celebraron al promediar el período de gobierno. La simultaneidad de las elecciones parlamentarias y presidencial en 2005 añade una nueva dimensión de incertidumbre para los líderes de la coalición concertacionista y para los potenciales candidatos al Congreso, ya que las negociaciones en torno a la composición de las listas para los comicios legislativos tendrán lugar antes de que se escoja el candidato presidencial de la Concertación por medio de una elección primaria. Frente a la perspectiva de un considerable riesgo electoral, y de una creciente inseguridad en cuanto a la posibilidad de obtener una compensación en caso de un resultado insatisfactorio, la pregunta es si los candidatos poderosos, y sus partidos asociados, seguirán dispuestos a mantener la coalición.

El mecanismo general de disolución de las coaliciones

Aun cuando el dilema que hemos descrito no es privativo de la Concertación, le es más agudamente aplicable porque se trata de la mayor coalición política de Chile (y por ende la que más tiene que perder en cuanto a su fondo de seguros), y también la más compleja, ya que abarca la mayor cantidad de actores políticos relevantes a nivel de subcoaliciones. Nuestra lógica se aplica a cualquier actor que se encuentre por debajo del nivel de la coalición —un candidato fuerte, un partido, o un bloque de

partidos al interior de una coalición más amplia— y que confíe en que si participa en una lista por su cuenta, teniendo el respaldo de un núcleo de adherentes, esa lista puede llegar por lo menos en segundo lugar y no ser doblada por ninguna otra. Si se cumplen dichas expectativas, este actor poderoso conseguirá un escaño en la circunscripción, incluso si se presenta como independiente. De manera que para un actor con estas características el beneficio potencial de pertenecer a una coalición no consiste en obtener un escaño parlamentario, sino más bien estriba en formar parte de una mayoría gobernante.

Sin embargo, integrar una coalición supone un riesgo electoral. La coordinación de una campaña junto con otros socios requiere llegar a acuerdos en las posiciones políticas, lo cual puede ocasionar un distanciamiento del núcleo de partidarios del actor. Si la coalición no dobla la votación de sus adversarios, y si los adherentes del actor se sienten suficientemente descontentos por las concesiones que exige la participación en una coalición, dicho actor puede incluso ser derrotado por su compañero de lista, perdiendo así el escaño parlamentario que hubiera podido ganar compitiendo solo. Cuando el riesgo puede ser indemnizado con la promesa de recibir un atractivo nombramiento en un cargo político, la peligrosa jugada que supone la participación en una coalición puede resultar atractiva para cualquier actor político que de otra manera se sentiría inclinado a postular por su cuenta en las elecciones parlamentarias. Nuestro argumento aquí es que dentro del actual panorama electoral chileno, a la Concertación le resultará cada vez más difícil lograr que esas promesas suenen creíbles, ya que, aun cuando la coalición se mantiene intacta, la posibilidad de que siga al mando del Poder Ejecutivo en el futuro es incierta en el mejor de los casos.

A decir verdad, cabría esperar que con una victoria presidencial prevista para el 2005, una dinámica similar facilitará la formación de listas en la Alianza por Chile. Si llega a la presidencia, la Alianza estará en condiciones de indemnizar a quienes hayan asumido riesgos en sus filas de candidatos parlamentarios. Los futuros subcampeones de la Alianza pueden abrigar la esperanza de recibir nombramientos compensatorios en altos cargos tras la elección, y estarán así más dispuestos a subordinar sus ambiciones electorales personales a los objetivos de la coalición.

Nuestros resultados también tienen aplicaciones teóricas transnacionales, especialmente en la medida en que los sistemas políticos de todo el mundo adoptan progresivamente sistemas electorales complejos en un intento por equilibrar la representación y la estabilidad políticas. A nuestro juicio resulta perfectamente razonable que en otros países existan sistemas de compensación para los candidatos derrotados que se expusieron a un

riesgo en beneficio de su coalición. Y puede que en otros países los sistemas de seguros sean menos decisivos para el mantenimiento de las coaliciones, ya que en todos los demás sistemas electorales con los que estamos familiarizados las economías de escala electorales son más directas que en el sistema binominal chileno. Es decir, cuando los distritos son más grandes (cuando en cada uno se eligen más de 2 representantes) la eficiencia con que los votos se convierten en escaños significa que disminuye la necesidad de inscribir a candidatos poderosos en las listas más marginales. Con todo, los estudios sobre el caso chileno sugieren la probabilidad de que los sistemas de compensación del tipo generado por el sistema electoral binominal se originen en otros sistemas electorales estratégicamente complejos. En cuanto a la teorización futura sobre los sistemas electorales y las instituciones informales, resulta fundamental analizar el grado en que la complejidad estratégica de los sistemas electorales puede conducir a la creación de instituciones informales para compensar la incertidumbre.

Para concluir, deseamos hacer hincapié en dos aspectos finales. Primero, no sugerimos que el sistema de subcampeones sea la única fuente de unidad, ni siquiera la más importante, de la Concertación. Sostenemos, sin embargo, que el sistema de subcampeones constituye un importante factor de cohesión que los analistas de la política chilena han pasado por alto, y que la desaparición de este sistema de recompensas, junto con otras tensiones que están emergiendo en el conglomerado, pueden acelerar considerablemente la disolución de la Concertación. En segundo lugar, no estamos propugnando el fin de la Concertación, ni tampoco estamos predispuestos a celebrar ese desenlace si se llegara a producir. A nuestro juicio, desde 1990 los gobiernos y las mayorías parlamentarias de la Concertación han proporcionado a Chile uno de los liderazgos políticos más lúcidos y moderados de la historia del país. Más bien presentamos nuestra predicción fatídica basándonos enteramente en lo que percibimos como un cambio fundamental en el panorama electoral chileno a contar del año 2000, y en lo que a nuestro parecer son los efectos de este cambio en la naturaleza estratégica de la formación y mantención de la coalición gobernante.

REFERENCIAS

- Carey, John M. "Parties and Coalitions in Chile in the 1990s". En Scott Morgenstern y Benito Nacif (eds.), *Legislative Politics in Latin America*. Nueva York: Cambridge University Press, 2002.
- Centro de Estudios Públicos. "Puntos Salientes: Encuesta Nacional de Opinión Pública". 2003. <http://www.cepchile.cl/>

- El Mercurio*. "Girardi Sentencia que ya No Existe la Concertación". Abril 13, 2003a. <http://diario.elmercurio.com>
- El Mercurio*. "Cargos Políticos se Reducirán a 760 de los actuales 3.500". Abril 25, 2003b. <http://diario.elmercurio.com>
- King, Gary; Michael Tomz y Jason Wittenberg. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation". *American Journal of Political Science* Vol. 44, N° 2 (2000), pp. 41-355. [Véase también <http://gking.harvard.edu/clarify/docs/clarify.html>]
- Latin America Adviser*. "Chilean Lawmakers Decry Leader's Statement on Ruling Coalition". Washington, D.C.: Inter-American Dialogue. Diciembre 11, 2002.
- Rabkin, Rhoda. "Redemocratization, Electoral Engineering, and Party Strategies in Chile, 1989-1995". *Comparative Political Studies*, 29 (1996).
- Siavelis, Peter. "The Hidden Logic of Candidate Selection for Chilean Parliamentary Elections". *Comparative Politics* Vol. 34 N° 4 (2002). □

BARRAS BRAVAS: TEORÍA ECONÓMICA Y FÚTBOL*

Enrique Gherzi

¿Cómo combatir a las “barras bravas”? ¿Qué sugiere al respecto el análisis económico del derecho? En las páginas que siguen, Enrique Gherzi invita a examinar el problema desde tres ópticas diferentes, cada una de las cuales sugiere un camino distinto para erradicar la violencia de los estadios: i) la responsabilidad civil extracontractual; ii) la responsabilidad penal-administrativa, y iii) los derechos de propiedad.

A juicio del autor, el hecho de que las dos primeras perspectivas terminen entregando respuestas insatisfactorias no es de extrañar, puesto que el origen del problema de la violencia en los estadios yace, primeramente, en la indefinición de los derechos de propiedad en el ámbito futbolístico. Luego, si se trata de garantizar la integridad personal y patrimonial del público asistente a los partidos de fútbol, debe proponerse una legislación que comience por reconstituir los derechos de propiedad en este deporte, destruidos por el monopolio de la FIFA, para que los incentivos estén donde deben estar y sean los propietarios del negocio los que se encarguen de “cuidarlo”.

ENRIQUE GHERZI. Abogado y escritor peruano. Profesor de la Universidad de Lima. Ex diputado por Lima. Francisco Marroquín Foundation Fellow. Miembro de la Sociedad Mont Pelerin. Académico Adjunto del Cato Institute. Coautor con Hernando de Soto y Mario Ghibellini de *El Otro Sendero*.

* Ponencia presentada en el VII Congreso Anual de la Asociación Latinoamericana y del Caribe de Derecho y Economía (ALACDE). Santiago de Chile, 13 y 14 de diciembre de 2002.

Debo a una pregunta de Alfredo Bullard y a la insistencia de Alberto Benegas Lynch (h) las siguientes reflexiones que quiero compartir con todos ustedes.

Supongamos por un momento que un legislador hipotético nos solicita que le preparemos un proyecto de ley para combatir el problema de la violencia en los estadios de fútbol. Para el efecto, nos pide recomendar una legislación que controle a las llamadas “barras bravas” y garantice la integridad personal y patrimonial de los asistentes al espectáculo, así como del público en general.

La presente ponencia busca ensayar una respuesta para nuestro hipotético legislador, desde la perspectiva del análisis económico del derecho¹.

Para proponer alternativas legislativas frente al problema de las llamadas “barras bravas” debemos partir de algunos supuestos de hecho muy importantes. En primer lugar, que están compuestas por hombres jóvenes, generalmente menores de edad. En segundo lugar, que las “barras bravas” rara vez se encuentran totalmente desligadas de los clubes.

Por el contrario, se ha sugerido muchas veces una conexión críticamente importante entre los clubes y estas organizaciones, llegando a pensarse que las estimulan de soslayo como una especie de fuerza de choque particular o que algunos dirigentes deportivos se sirven de ellas para satisfacer sus expectativas. En tercer lugar, que se sabe poco de su organización interna, salvo que generalmente agreden a sus víctimas en los estadios o en sus cercanías, los días de partido de fútbol o en sus inmediaciones. Se trata, pues, de organizaciones con una acción previsible y temporal. No se conocen como entidades permanentes, aunque muy recientemente se han registrado agresiones con saldos mortales como actos aislados y sin necesaria proximidad con los partidos programados.

En base a tales consideraciones, esta ponencia analizará las posibilidades legislativas desde la perspectiva de la responsabilidad civil extracontractual, de la responsabilidad penal-administrativa y de los derechos de propiedad que, como veremos posteriormente, es nuestra opción preferida.

¹ Si se quiere información introductoria sobre esta disciplina, se sugiere consultar Richard Posner, *Economic Analysis of Law* (1977); Hans-Bernd Schafer y Claus Ott, *Manual de Análisis Económico del Derecho Civil* (1991); Juan Torres López, *Análisis Económico del Derecho* (1987); y Mitchell Polinsky, *Introducción al Análisis Económico del Derecho* (1985).

1. Responsabilidad civil extracontractual

Analizaremos las siguientes opciones de responsabilidad civil, con especial atención a si podrán reducir los daños causados por las “barras bravas” y/o compensar a las víctimas:

1.1. Seguro mutual

Tomando en consideración que las “barras bravas” pueden entenderse como un riesgo creado por los espectáculos de fútbol, podría proponerse un sistema de responsabilidad objetiva consistente en que todos los asistentes a tales espectáculos paguen un sobreprecio en su entrada para crear un fondo que se administraría como una suerte de seguro mutual. Este fondo se destinaría a cubrir los daños causados por los hinchas a cualquier persona o patrimonio.

La justificación de una propuesta así podría encontrarse en que la agresividad es colectiva, de manera que las “barras bravas” son un riesgo creado por el fútbol que todos los aficionados están en la obligación de compensar a la sociedad. Su ventaja principal es que, al ser de fácil recaudación, permitiría asegurar cabalmente las compensaciones para las víctimas, especialmente si tenemos en cuenta que los daños por vandalismo pueden ser fácilmente más altos que lo que los patrimonios individuales de los vándalos pueden satisfacer. Otra ventaja conexa es que, por lo general, los peores agraviados son los propios asistentes a los partidos, de manera que podría pensarse que es justo asegurar su propia imprevisión al asistir a espectáculos potencialmente riesgosos sin ninguna cobertura.

No obstante todo lo dicho, una propuesta de responsabilidad objetiva como la sugerida también tiene desventajas. La principal es el llamado “peligro moral”. En efecto, si se cubren todos los eventos que pudieran producir los vándalos a través de un seguro mutual, efectivamente no existirán daños que queden insatisfechos, pero no habrá tampoco ningún incentivo para que los asistentes al fútbol desarrollen comportamientos más deseables. En otras palabras, como todos los asistentes pagarían los daños de las “barras bravas”, no habría ningún estímulo para que la gente se porte bien en los estadios.

Por el contrario, podrían inclusive producirse respuestas no cooperativas de algunos asiduos asistentes que, cansados de pagar a lo largo del tiempo su cuota al seguro mutual sin tener beneficio perceptible de la misma, decidieran portarse mal también ellos para que la transferencia de

recursos que han hecho no quedase totalmente falta de amortización por su parte.

Otra desventaja adicional es que debería crearse algún sistema de administración para este seguro. De suyo, este sistema podría ser público o privado. Una administración pública implicaría una suerte de estatización del fondo de seguro que introduciría las distorsiones burocráticas habituales en el sistema; el cual en este supuesto pasaría a parecerse más a lo que en teoría se conoce como una contribución de Pigou².

Por su parte, una administración privada, aunque consumiría también recursos, podría hacerse a través de mecanismos muy simples, dando en concesión o licitando el servicio. Empero, en ambos casos la necesidad de administrar el sistema especialmente quizás involucre una distracción importante de recursos del fondo de seguro, pudiendo llegarse incluso a situaciones de insolvencia conocidas en la experiencia de la mal llamada “seguridad social”.

En consecuencia, creemos que esta alternativa es insatisfactoria, pues, aunque fuese capaz de compensar a las víctimas, no va a reducir los daños ocasionados por las “barras bravas”, pudiendo generar además comportamientos imprevisiblemente riesgosos y un enorme desperdicio de recursos. La sola idea de pensar en fanáticos con seguro resulta realmente aterradora.

1.2. Otros seguros

En teoría podría argumentarse que es posible corregir los defectos del sistema de seguro mutual, por definición un seguro imperfecto, mediante sistemas comerciales de aseguramiento. En este caso, la legislación propuesta sería también de responsabilidad objetiva y se limitaría a establecer que todo asistente a los estadios debe tener un seguro obligatorio, permitiéndose que participen todas las compañías de seguros del mercado.

Cada empresa tendría que calcular el costo de las primas. No habría un pago igual por asistente, sino que éste reflejaría tanto los costos probables de un siniestro como la peligrosidad y previsión de cada uno. De esta forma —podría argumentarse— no sólo quedaría satisfecha la expectativa de las compensaciones, sino que se dejaría a salvo el llamado peligro moral, ya que vía prima se colocarían estímulos para que la gente reduzca los daños y tenga un mejor comportamiento.

² Cf. Hans-Bernd Schafer y Claus Ott, *Manual de Análisis Económico del Derecho Civil* (1991), p. 123.

No obstante, esta corrección al sistema del seguro mutual propuesto no nos parece satisfactoria. Principalmente porque es imposible. Una entrada a un estadio de fútbol es un servicio de muy pequeño valor económico como para soportar todos los costos que implicaría un sistema de seguros perfecto. Terminaría encareciéndose enormemente el espectáculo con las pesquisas necesarias para la determinación de las primas, registros de siniestros, archivos individuales, exigencias de previsión, etc. En este supuesto, dado el carácter instrumental del derecho³, se produciría un desplazamiento de la demanda de espectáculos deportivos de la formalidad hacia la informalidad, de tal manera que la gente que simplemente no pudiera pagar el costo del seguro obligatorio en los estadios, acudiría a partidos de fútbol ilegales, donde por cierto se reproducirían los problemas de las “barras bravas” sin limitación alguna.

Dado que la ley es independiente de las preferencias de la población y constituye sólo un medio a través del cual la gente las satisface, una de las más importantes limitaciones para una legislación contra las “barras bravas” está dada por el nivel de costo que la gente estaría dispuesta a soportar por un partido de fútbol. Si se le cobra demasiado a la gente por un partido, el fútbol no desaparece, simplemente se traslada, con “barras bravas” y todo.

1.3. Responsabilidad de los clubes

Frente al hecho de que la responsabilidad objetiva con un sistema de seguro imperfecto asegura la compensación pero no reduce los daños y con un sistema de seguro perfecto resulta tan onerosa que puede crear un mercado negro de fútbol, podría ensayarse la responsabilidad por culpa. Sin embargo, parece obvio que una responsabilidad por culpa tradicional en cabeza de los fanáticos resulta ineficiente, dados sus altos costos de identificación y procesamiento, por lo que vamos a abstenernos de analizarla. Además, por existir ya prácticamente en toda la legislación occidental, no requeriría de modificación legal alguna para satisfacer a nuestro hipotético legislador. En su lugar podríamos sugerir otra alternativa dentro del sistema de culpa: hacer a los clubes responsables por los daños causados por las “barras bravas”.

En este caso lo que se estaría haciendo es, dentro de una responsabilidad por culpa, trasladar el pago de la compensación del culpable individual a un hipotético culpable colectivo o, si se quiere, a un culpable intelec-

³ Cf. Enrique Ghersi, “El Costo de la Legalidad” (1988), pp. 88-89.

tual. Aunque desde una perspectiva tradicional puede sonar arbitrario, resulta una opción sugerente, sobre todo porque sería posible elaborar un argumento justificatorio a partir de la responsabilidad vicaria⁴. Si definimos ésta como aquella responsabilidad que corresponde a quienes tienen a sus órdenes a terceros por los actos que ellos lleven a cabo en el desempeño de su función, podría sostenerse que, dada la relación entre los clubes y las barras, la naturaleza jurídica correspondiente a ella es la de la responsabilidad vicaria, de suerte que corresponde pagar toda compensación por los excesos de los barristas al propio club deportivo.

En este caso la legislación a proponerse podría establecer que los clubes paguen los daños causados por los vándalos a personas y propiedades. A esta fórmula simple convendría, sin embargo, agregarle un agravante, consistente en que frente a circunstancias especialmente dañosas o peligrosas el castigo al club no solamente sea en dinero, sino también en pérdida de puntos en el campeonato. De esta forma se buscaría corregir una probable distorsión. A saber, que las “barras bravas” puedan no tener ningún escrúpulo en que el club pierda plata, pero sí frente a la posibilidad de que su equipo pueda quedar perjudicado en la competencia deportiva. Sería una especie de castigo al fanatismo.

Esta alternativa, de hecho, no es original. Hace algunos meses fue introducida en ciertos países como una fórmula presuntamente de mercado dirigida a forzar a los clubes a controlar a sus hinchas y castigar a los fanáticos directamente donde les duele; es decir, en los resultados de los equipos de su preferencia.

En favor de esta fórmula se puede argumentar que los clubes no son inocentes frente a las “barras bravas”. Como explicamos al principio, diversas investigaciones periodísticas sugieren que, por lo menos en una etapa incipiente, son ellos los que las organizan y mantienen, facilitándoles hasta las entradas y pasajes a las diferentes ciudades en que los equipos tienen que presentarse. Luego, si los clubes quieren tener hinchas leales, podría decirse que ellos son los llamados a cubrir los costos que ocasionen los extravíos de tales lealtades.

Simétricamente, podría decirse que si los clubes no son inocentes respecto a las “barras bravas”, son ellos quienes mejor pueden controlarlas al tener que soportar los costos económicos y deportivos de los excesos que ellas produzcan. Estando en contacto con los barristas y dependiendo éstos en mucho del apoyo de los clubes para poder seguir a los equipos de estadio en estadio y de ciudad en ciudad, son esas entidades las que están más cerca de influir positivamente sobre ellos.

⁴ Cf. Fernando De Trazegnies, *La Responsabilidad Extracontractual* (1988), tomo I, p. 467.

Una segunda argumentación en favor de esta alternativa sería que, por lo general, los clubes son más solventes que sus hinchas, de suerte que haciendo responsables a aquéllos, se asegura convenientemente el pago de las compensaciones probables que se deriven de los actos vandálicos.

Entonces, podría concluirse que esta alternativa satisface el doble propósito de reducir la cantidad de daños y asegurar las compensaciones para las víctimas. Sin embargo, a nuestro entender, existen algunas objeciones importantes que hacer a una propuesta de legislación de esta naturaleza.

La primera objeción es que los sistemas de responsabilidad por culpa tienden a la ineficiencia porque resultan excesivamente costosos para obtener resultados satisfactorios⁵.

La segunda objeción está relacionada con la clásica noción de culpa. ¿Es justo que un tercero tenga que cargar con los costos de los extravíos de otros, aunque sean sus admiradores? ¿No sería como hacer a Madonna responsable por los actos de sus fans o a la Iglesia Católica por los destrozos de sus fieles? Ciertamente que la noción de culpa aparece totalmente reñida, en casos de este tipo, con la justicia y el objetivo de eficiencia que se busca.

Una tercera objeción está relacionada con las señales equivocadas que crearía esa legislación, pues importaría una suerte de sanción a la popularidad y el éxito. Un castigo para el que se destaca y es capaz de tener arrastre entre la población. Una suerte de segregación al que ha sido suficientemente hábil como para suscitar la fe inquebrantable de los demás. Las distorsiones resultantes serían tanto más imprevisibles si se adoptara una legislación semejante sólo para el deporte, cuando es perfectamente posible que en otras áreas como la política, el arte o la religión puedan presentarse igualmente.

La cuarta objeción a esta propuesta es que supone una entidad juzgadora que debe imponer las sanciones. Esta entidad debe ser obviamente especializada y parece necesario que sea completamente independiente de las actuales instituciones deportivas, a fin de garantizar el grado de autonomía necesario para imponer las sanciones. El sistema dependerá, en realidad, de que una entidad de este tipo sea capaz de actuar con la suficiente justicia, oportunidad y firmeza para que los incentivos que teóricamente se quieren introducir realmente funcionen. La amplia discrecionalidad que debe tener es la clave de su éxito. En el fondo, pues, este sistema se basa en la existencia de unos jueces justos —en el sentido amplio del término— que sean capaces de ponerlo en funcionamiento. Esto, a nuestro entender,

⁵ Cf. Paloma Durán Lalaguna, *Una Aproximación al Análisis Económico del Derecho* (1992), pp. 126-127.

es insatisfactorio, pues la gente justa y proba constituye un bien socialmente muy escaso que seguramente no estará disponible para el fútbol, si de hecho no lo está para cosas más trascendentes como la propia administración de justicia⁶.

La última objeción que encontramos, a nuestro juicio tal vez la más importante, es que una legislación de este tipo puede generar reacciones imprevisibles entre los fanáticos. La más obvia es que si las “barras bravas” saben que el que paga es el club, en términos económicos y deportivos, tienen una conciencia muy clara de que esto perjudica. Luego, los barristas pueden disfrazarse como de otro club y provocar graves incidentes a fin de perjudicar al rival y obtener por esa vía inclusive los resultados deportivos que en la cancha no pudieron conseguir sus jugadores.

Por ejemplo, en un campeonato de fútbol muy reñido en que el título se obtiene con un punto, bien podría pasar por la cabeza de las barras o de los clubes que, como hemos dicho, no parecen inocentes frente de ellas, provocar un incidente haciendo pasar algunos fanáticos como hinchas del otro equipo a fin de lograr que se imponga al enemigo deportivo sanciones injustas. Se estaría creando sin querer una suerte de rentismo negativo derivado de la utilización de la ley por los operadores del derecho.

Este problema está también relacionado con el carácter instrumental del derecho que explicamos en nuestras objeciones a las diferentes modalidades de seguros propuestas para enfrentar a las “barras bravas”. Esta instrumentalidad significa que la ley es inelástica; es decir, que la gente busca satisfacer sus preferencias y deseos cuando cumple con la ley y no las preferencias ni los deseos de los legisladores. Luego, si las barras pueden utilizar la legislación para seguir causando siniestros, en este caso a través de lo que en teoría podríamos llamar renta negativa, lo van a hacer con independencia de cuáles han sido los objetivos de las normas así propuestas.

Aunque en apariencia podría parecer que este sistema puede funcionar, encierra peligros muy grandes. El principal es que su eficiencia no depende del sistema mismo sino del orden institucional vigente en el país, debido a que se basa esencialmente en la coacción y la coerción, así como en la total prescindencia de las reacciones de los fanáticos y de los intereses de los clubes, de suerte que podría estarse generando un sistema de rentas deportivas por manipulación de las normas.

La principal de tales rentas sería el desplazamiento de la competencia del puro ámbito deportivo al político e institucional, en el cual las

⁶ Cf. Andrés Roemer, *Introducción al Análisis Económico del Derecho* (1994), pp. 26-35.

consecuencias finales podrían consistir en que un campeonato ya no se defina en el estadio sino en el Congreso que, como la propia historia enseña, no es sino una de sus variantes acaso más rudimentarias.

Estos problemas derivan de no haber entendido que el origen de las “barras bravas” no es la conspiración entre clubes y fanáticos que se deba castigar sino, como veremos posteriormente, la ausencia de derechos de propiedad claramente establecidos en el sistema futbolístico internacional.

1.4 . Responsabilidad de clubes y socios

Como una variante de la propuesta anterior podría sugerirse a nuestro hipotético legislador combinar la responsabilidad de los clubes con la responsabilidad de sus socios, de tal manera que las compensaciones económicas se repartan entre ellos según una proporción a señalar.

La racionalidad de esta propuesta reside en que si los clubes no son inocentes frente a las barras, es porque las dirigencias de esos clubes así lo quieren. Luego, a fin de controlar a las dirigencias, se debe trasladar a los socios una parte del costo a efectos de que éstos puedan tomarles cuentas a sus dirigentes y eventualmente deponerlos si no ponen coto a los desmanes de las barras. Si sólo paga el club es probable que muchos ni siquiera se enteren de lo que está pasando. Si, en cambio, todos los socios tienen que pagar una parte, existirían estímulos para un control descentralizado de la violencia.

Las objeciones a esta variante son, básicamente, las mismas que ya hicimos en el numeral anterior. No parece una alternativa recomendable.

2. Responsabilidad penal-administrativa

Dado el carácter de agresión y no de accidente que tiene la actividad de las “barras bravas”, es perfectamente posible sostener que las medidas más adecuadas para combatirlas no deben venir de la perspectiva de la responsabilidad civil, sino del derecho público. Esto se encontraría doctrinariamente justificado en que el vándalo actúa dolosamente, por lo que se puede tipificar su conducta válidamente como criminal, reforzando además el control *ex ante* de la acción de las barras que los mecanismos de responsabilidad civil dejan librados al resultado de los incentivos resultantes de sus controles *ex post*.

A este nivel podrían sugerirse las siguientes iniciativas legislativas:

2.1. Tipificación de un delito especial

Desde una perspectiva estrictamente penal, es posible proponer la reforma del Código correspondiente para tipificar, como un agravante de los delitos de lesiones y daños, las acciones de las “barras bravas”. Esta reforma podría incrementar las penas con especial severidad, privando a los encausados de algunos beneficios como la libertad provisional y la caución, pero debería establecer también penas anexas que llevarán a algún tipo de inhabilitación al agresor.

En la antigüedad, a los autores de ciertas agresiones se les marcaba corporalmente para el escarnio público. Como esto se encuentra reñido con los valores predominantes en la actualidad y podría resultar además paradójico entre los vándalos que, de repente, estarían encantados de exhibir sus marcas como trofeos de guerra, la mejor forma de crear penas conexas que acarreen inhabilitación es la de establecer, probablemente para un agresor reiterante, algún tipo de marca en su documento de identidad que, mediante mecanismos indelebles, informara a toda persona que está frente a un vándalo de las “barras bravas”. Esto podría tener efectos disuasivos especialmente frente a las oportunidades de empleo, porque resultaría obvio que ningún patrón tendría interés en contratar a gente con antecedentes comprobados de violencia. Aunque podría ocurrir también que empleadores fanáticos del mismo club encontrasen este elemento como una forma fácil de seleccionar personal de confianza.

Una variante más en esta alternativa consistiría en agregar a la tipificación del delito especial y al incremento de las penas, el concepto del tercero civil⁷, con lo cual se estaría elaborando una alternativa combinada con las formas de responsabilidad civil anteriormente analizadas.

La doctrina penal reconoce en casos culposos y, en general, de delitos cometidos por personas bajo el mando de otras, que se puede involucrar en el proceso a un tercero llamado a pagar el monto de la llamada reparación civil. En este caso podría sugerirse una combinación de la represión penal con la responsabilidad de los clubes o con la responsabilidad de los clubes y de los socios que vimos en los numerales 1.3 y 1.4.

La desventaja principal de estas proposiciones estriba en la limitación de la ley penal. En primer lugar, porque se tiene que identificar personalmente al autor de los desmanes; cosa no siempre sencilla en los actos de masas. En segundo lugar, porque hay presunción de inocencia y se debe probar específicamente la relación causal entre los actos realizados indivi-

⁷ Cf. María Alejandra Acero Colmenares y Juan Alberto Castro Flórez, *El Tercero Civilmente Responsable en el Proceso Penal* (1989), p. 5.

dualmente y el resultado dañoso. En tercer lugar, porque debe probarse el dolo, pues sin la evidencia del elemento intencional no es posible aplicar un castigo penal.

Todo ello puede conducir a que la represión penal quede neutralizada por las dificultades de llevarse a efecto, convirtiendo a la disuasión esperada en un mero saludo a la bandera. Coadyuvaría a tal resultado el hecho adicional de que la justicia penal es la que peor funciona y que, estando las cárceles llenas con delincuentes más peligrosos, no es razonable esperar condenas de prisión efectiva por estos delitos, simplemente porque no habría dónde poner a los sentenciados.

En lo que se refiere a la introducción de la figura del tercero civil en estos casos, reiteramos nuestras objeciones anteriormente expuestas en el capítulo 1.

2.2. Registro de infractores

A caballo entre lo penal y lo administrativo, se podría aconsejar la constitución de un registro de vándalos, de tal manera que estén prohibidos de entrar a los estadios los incluidos en él. Estas prohibiciones podrían graduarse en base a las ofensas, a su gravedad y a la reiteración, yendo desde restricciones temporales hasta prohibiciones absolutas.

En algunos países desarrollados como Gran Bretaña, la policía ha introducido ya este tipo de controles en las puertas de los estadios. Se basan en la pura y simple interdicción y en la conveniente administración de una información que ya está disponible en comisarías.

Empero, esta sugerencia exigiría superar dos dificultades. En primer lugar, el procedimiento mediante el cual se aplica la sanción al infractor. En segundo lugar, el procedimiento de control para llevar a cabo la sanción. El uno puede ser atribución policial, en el estilo de las infracciones de tránsito. El otro exige establecer un control computarizado en el acceso a los estadios a fin de que no penetren los prontuariados, en el estilo de los controles migratorios. Lamentablemente, en ambos casos el principal limitante es la corrupción.

En este contexto, si hacemos un paralelo entre lo que ocurre con las infracciones de tránsito y los controles migratorios, no existe razón alguna para creer que este sistema pueda funcionar realmente. Además, basta recordar que la oferta se desplaza por el carácter instrumental del derecho del que ya hemos hablado. En otras palabras, esto significa que, de existir controles muy eficientes que detectarían a los vándalos, lo previsible es que

aparezcan partidos de fútbol ilegales en los que no haya control alguno. Este surgimiento de mercados negros deportivos, del que ya hemos hablado anteriormente, no es más que la clásica consecuencia del desplazamiento de oferta en presencia de regulaciones ineficientes⁸.

2.3. Otras restricciones

Es posible aconsejar otras restricciones de índole administrativa para controlar la acción de los vándalos. Si, como hemos dicho, éstos son por lo general menores de edad, se puede establecer que los partidos de fútbol sean sólo espectáculos para mayores. De esta forma, se disminuiría la exposición de lo que es el grupo socialmente más exaltado y belicoso.

Es también posible aconsejar restricciones no sólo de edad sino de sexo. Visto el hecho de que los integrantes de las “barras bravas” son hombres, es posible establecer que solamente las mujeres de cualquier edad y los hombres mayores de edad puedan asistir a los estadios.

El problema principal con estas restricciones es que parecen poco capaces de conseguir resultados por ellas mismas, siendo probablemente de índole complementaria, amén de que alguien podría considerarlas como odiosamente discriminatorias. Otras restricciones como la prohibición del alcohol en los estadios parecen obvias y nos relevan de mayor comentario.

3. El problema de los derechos de propiedad

Mientras hemos venido haciendo el análisis de este problema desde la perspectiva de la responsabilidad civil y de la responsabilidad penal-administrativa nos fue surgiendo la duda de si la falta de respuestas plenamente satisfactorias no es consecuencia, en realidad, de que el origen del problema no estriba en nada de esto sino en una indefinición de los derechos de propiedad en el ámbito futbolístico.

Una serie de apreciaciones nos persuaden de que es así. En primer lugar, nadie ha oído hablar de “barras bravas” en otros deportes que no sean el fútbol. Por ejemplo, no hay violencia en los hipódromos, pese a que los riesgos económicos son ahí mucho mayores. Excluida la violencia del deporte mismo, ni en el box ni en las artes marciales existen cosas parecidas a las “barras bravas”. No hay violencia en el golf ni tampoco en el

⁸ Cf. Svetozar Pejovich, *Fundamentos de Economía. Un Enfoque Basado en los Derechos de Propiedad* (1979), apéndice X.1, pp. 154-161.

vóleibol. Esto sugiere poderosamente que algo pasa en el fútbol que no ocurre en los demás deportes y que puede estar relacionado con sus reglas internas.

En segundo lugar, se debe descartar el que sea un problema de deportes masivos, en realidad. Si examinamos algunos de éstos, sin embargo, tampoco encontramos cosas tales como las “barras bravas”. En el básquetbol no las hay. Tampoco en el béisbol, en el hockey y, ni siquiera, en el fútbol americano.

Luego, algo debe suceder con las reglas internas. Efectivamente, mientras que en el béisbol, el básquetbol, el hockey y el fútbol americano los derechos de propiedad están claramente establecidos y cada equipo, cada estadio y hasta el deporte mismo tiene un dueño, en nuestro fútbol existe una total indefinición en ese sentido. No sólo porque muchos de los estadios son de propiedad pública, sino principalmente porque los clubes y el deporte mismo no son de propiedad de nadie. Al ser esto así, no hay el más mínimo interés en cuidar del negocio a largo plazo, sino de disfrutarlo en el corto y al estilo político, de suerte que queda establecida la base sobre la cual la grandeza del dirigente puede construirse sobre la adoración de una hinchada fanática que es, también, una fuente de su riqueza y poder.

Lamentablemente, éste no es sólo un problema nacional. El fútbol se halla dominado por la FIFA a nivel internacional y esta institución, superestado intervencionista que lo regula todo y que recibe una renta de todo, conforma una especie de monopolio que controla totalmente la forma en que se desarrolla y organiza la actividad. En algunos casos llega a actuar hasta como monopsonio, puesto que a través de los campeonatos se instituye en el principal consumidor de ese deporte.

En realidad el sistema impuesto por la FIFA se basa en relaciones personales y no en relaciones de propiedad. Diferentes caudillos se entronizan en la dirección del organismo y por cooptación designan a los encargados de dirigirlo a nivel internacional. No hay relación dominal alguna. Existe vínculo de lealtad, camarilla, afinidad. Una organización de este tipo destruye los estímulos de largo plazo y propende a que se preste poca atención al mantenimiento del negocio o su difusión, pues resulta más importante atender al jefe, al colega o la estrella en ascenso de la dirigencia mundial.

Es cierto que en algunos países los clubes tienen formalmente un dueño y aun los estadios; cosa que ni siquiera ocurre en los nuestros. Pero no lo es menos que llamarse dueño de un club, en un contexto de monopolio privado consagrado por las legislaciones deportivas de prácticamente todos los países del orbe, no significa nada. En el fútbol la FIFA es dueña de todo, pero nadie es dueño de la FIFA.

No interesa saber quién es propietario de las instalaciones o de los uniformes deportivos. Lo importante es saber quién es el dueño del negocio. Como en la educación, no interesa que los ladrillos y las carpetas sean de propiedad individual para que exista educación privada.

Lo que interesa es que el currículum sea definido por los dueños formales de las escuelas. Utilizándose una paráfrasis informática, podría decirse que la propiedad del fútbol no es un problema de hardware, es un problema de software.

Mucho se ha dicho, por ejemplo, que el básquetbol, el fútbol americano o el hockey, en los Estados Unidos, se encuentran dominados por la mafia. Puede que ello sea cierto. Para el caso no interesa, pues estando claramente establecidos los derechos de propiedad, nadie permite que haya violencia en los estadios, porque ello destruye la afluencia de público y daña las posibilidades de desarrollo del negocio. Ningún propietario auténtico va a consentir que se produzcan actos como los de las “barras bravas” que van a repercutir directamente en contra de sus propios intereses.

Luego, si se trata de combatir este problema, la mejor legislación que puede proponerse es aquella que reconstituya los derechos de propiedad en este deporte, para que los incentivos estén donde deben y sean los propietarios del negocio los que se encarguen de cuidarlo, reduciendo la cantidad de daños, compensando a las víctimas y combatiendo hasta desaparecer a las “barras bravas”. En nuestro concepto, la ventaja principal de este sistema es que coloca los incentivos en su sitio, no demanda gasto público y tiende a mediano plazo a producir el resultado más próximo al óptimo de Pareto⁹.

Esta situación, en donde no es posible desarrollar una actividad futbolística libre, ha derivado en una destrucción de los derechos de propiedad en este deporte.

En términos estrictos, las “barras bravas” son consecuencia de un ambiente institucional en el que no existe derecho de propiedad. La literatura especializada llama a este fenómeno *free rider*, que podría traducirse como polizón o —más criollamente— como “gorreón”.

North y Thomas lo definen como “aquella situación que se presenta en ciertos bienes públicos de cuyo consumo nadie puede quedar excluido y por el cual cada individuo busca beneficiarse de las contribuciones de los demás”¹⁰.

⁹ Cf. Pedro Schwartz y Alfonso Carbajo Isla, “Teoría Económica de los Derechos de Apropiación” (1980), p. 134.

¹⁰ Douglass C. North y Robert Paul Thomas, *El Nacimiento del Mundo Occidental* (1978), p. 11.

El *free rider* o fenómeno del polizón significa que si no somos dueños de las cosas, sino que éstas son de todos, siempre habrá gente que se beneficie de la conducta diligente de otras sin tener que hacer nada para el efecto. Se trata de una actitud no cooperativa típica en que los estímulos para comportarse como es debido se reducen porque otros tienden a comportarse así.

Por lo general, este concepto es utilizado para explicar por qué unos pagan impuestos y otros no, o por qué unos limpian las calles y otros no. Somos de la opinión de que es, asimismo, pertinente para esclarecer el tema de las “barras bravas”.

En general, el fútbol está contaminado por el fenómeno del polizón porque no tiene dueño. Barristas y dirigentes no son más que polizones dispuestos a capturar un deporte del que no son dueños pero en el que se beneficiarán de la energía, el talento y la pasión de atletas y aficionados.

Podría argumentarse en contrario, que siendo la estructura institucional del fútbol consecuencia espontánea de la cooperación voluntaria de clubes individuos en el mundo, este carácter cataláctico hace que el resultado sea el mejor de los posibles.

Tal observación, sin embargo, es inexacta, por lo menos, por las siguientes razones:

1. Es harto discutible que la estructura institucional del fútbol sea un producto cataláctico. En su origen sin duda lo fue, por cuanto las reglas y la organización se originaron espontánea y competitivamente. Desde hace décadas, sin embargo, esto no es así. Muchas veces a instigación de la propia FIFA y las federaciones nacionales, los gobiernos se han estado involucrando en la elaboración de las reglas y la organización de este deporte. Inclusive, podría identificarse una curiosa tendencia mercantilista según la cual la FIFA y sus federaciones integrantes estimulan a que los diferentes Estados adopten como legislación nacional las normas y sugerencias que la propia FIFA propone. Esta influencia llega a tal extremo de que si por ventura algún gobierno osa dictar una legislación que no es del agrado de la FIFA, la sola amenaza de ésta de desafiliar al respectivo seleccionado nacional basta para conseguir que el gobierno aludido retroceda.
2. Un resultado cataláctico puede ser insuficiente y defectuoso. En efecto, existiendo costos de transacción, es la estructura institucional la que puede acercar o alejar una conducta humana del óptimo económico. Ya Henri Lepage ha acreditado, por ejemplo, que las socie-

dades de personas son ineficientes, pues tienen altos costos de transacción y tienden a una elevada politización¹¹. En *El Otro Sendero*, por lo demás, encontramos esa misma característica en las organizaciones informales, con el rasgo adicional de que se propende a la violencia por la falta de definición en los derechos de propiedad¹². Ello hace que, en general, las sociedades de personas —cooperativas, asociaciones, clubes, etc.— tiendan a convertirse en organizaciones políticas. Como tales, la competencia se desplaza del ámbito económico al político y el resultado deja de ser cataláctico, toda vez que lo que se busca es el favor de un gobierno, no el del consumidor.

Nada en esta discusión pretende sugerir que el fútbol está condenado a promover violencia ni mucho menos. Lo que se quiere sugerir es que un determinado estadio de evolución institucional, las reglas y la organización del fútbol, dada la indefinición de los derechos de propiedad que le es característica, provoca violencia como la de las barras bravas. Es probable, además, que la propia evolución institucional proporcione los mecanismos de corrección de este fenómeno. Por ejemplo, los clubes de fútbol han comenzado a convertirse en sociedades de capital y a dejar de ser sociedades de personas. De hecho, el Manchester United y La Juventus están listados en bolsa. Si como es de suponer este proceso cambiara y se produce espontáneamente una evolución institucional tal que empezaran a definirse derechos de propiedad en este deporte, es posible que se espere una desaparición correspondiente de la violencia.

Luego, la violencia en el fútbol no es la causa sino el efecto del problema: la falta de propiedad. Para enfrentarla, consiguientemente, hay que reconstituir conductas cooperativas que lograrán como Tullock dijo, “el bienestar de la sociedad para la mayoría sólo si hay algún beneficio privado para nosotros al actuar en este sentido”¹³.

Las “barras bravas”, los hooligans, la violencia que devastan el fútbol en el mundo son la consecuencia de la destrucción de los derechos de propiedad en ese deporte por el monopolio de la FIFA. Son, si se quiere, un típico ejemplo de cómo todos tratan de vivir a costa de los demás.

¹¹ Henri Lepage, *Autogestión y Capitalismo* (1978).

¹² Hernando de Soto, Enrique Ghersi y Mario Ghibellini, *El Otro Sendero* (1986).

¹³ Gordon Tullock, *Los Motivos del Voto* (1979), p. XXXV.

REFERENCIAS BIBLIOGRÁFICAS

- Acero Colmenares, María Alejandra y Juan Alberto Castro Flórez. *El Tercero Civilmente Responsable en el Proceso Penal*. Bogotá: Pontificia Universidad Javeriana, Facultad de Ciencias Jurídicas y Socioeconómicas, 1989.
- De Trazegnies, Fernando. *La Responsabilidad Extracontractual*. Vol. IV (2 tomos). Biblioteca para Leer el Código Civil. Fono Editorial PCUP (Pontificia Universidad Católica de Perú) Lima, 1988.
- Durán Lalaguna, Paloma. *Una Aproximación al Análisis Económico del Derecho*. Granada: Comares Editorial, 1992.
- Gheri, Enrique. "El Costo de la Legalidad". *Estudios Públicos*, N° 30 (otoño 1988).
- Lepage, Henri. "Autogestión y Capitalismo". Madrid: APD, 1978.
- North, Douglass C. y Robert Paul Thomas. *El Nacimiento del Mundo Occidental*. Madrid: Siglo XXI de España Ed. S.A., 1978.
- Pejovich, Svetozar. *Fundamentos de Economía. Un Enfoque Basado en los Derechos de Propiedad*. México: Fondo de Cultura Económica/Serie Economía, 1979.
- Polinsky, Mitchell. *Introducción al Análisis Económico del Derecho*. Barcelona: Ariel, 1985.
- Posner, Richard. *Economic Analysis of Law*. Little Brown & Co, 1977.
- Roemer, Andrés. *Introducción al Análisis Económico del Derecho*. México: Instituto Tecnológico Autónomo de México-Sociedad Mexicana de Geografía y Estadística-Fondo de Cultura Económica, 1994.
- Schafer, Hans-Bernd y Claus Ott. *Manual de Análisis Económico del Derecho Civil*. Madrid: Tecnos, 1991.
- Schwartz, Pedro y Alfonso Carbajo Isla. "Teoría Económica de los Derechos de Apropiación". En *La Nueva Economía en Francia y España*. Madrid: Colección Fórum Universidad-Empresa, 1980.
- Soto, Hernando de, Enrique Gheri y Mario Ghibellini. *El Otro Sendero*. Lima: Editorial El Barranco, 1986.
- Torres López, Juan. *Análisis Económico del Derecho*. Madrid: Tecnos, 1987.
- Tullock, Gordon. *Los Motivos del Voto*. Madrid: Espasa Calpe, 1979. □

EXIGENCIAS PARA LA CONSTRUCCIÓN DE UNA PRUEBA DE SELECCIÓN A LA UNIVERSIDAD

Bárbara Eyzaguirre

Con el objeto de regular el proceso de construcción de pruebas y garantizar un uso adecuado de sus resultados, se han desarrollado estándares que prescriben lo que se considera una buena práctica en la elaboración y el uso de instrumentos de evaluación. Este artículo se propone dar a conocer esos estándares, y, a la luz de ellos, evaluar la prueba de admisión a la educación superior utilizada en Chile hasta el 2002 (PAA), así como el proyecto SIES y la prueba PSU, actualmente en preparación, que se aplicará a fines del 2003.

Del análisis se concluye que el procedimiento seguido en la construcción de la nueva prueba de admisión (PSU) no ha sido el regular, y que son mínimas las posibilidades de que ésta llegue a cumplir, antes de su aplicación, con los estándares requeridos.

Por ello, se señala finalmente, la prueba de admisión 2003 no debería constituirse en un marco forzado para las evaluaciones de los años subsiguientes, sino que se la debería considerar como un antecedente más en el estudio de la batería idónea para seleccionar a los mejores alumnos para la universidad. En este sentido, se sugiere

BÁRBARA EYZAGUIRRE. Psicóloga educacional especializada en desarrollo cognitivo, con experiencia en programas de mejoramiento de la calidad de la educación en sectores de escasos recursos. Fundadora y asesora pedagógica de la Escuela San Joaquín (Renca), perteneciente a la Fundación Marcelo Astoreca. Investigadora del Centro de Estudios Públicos.

Se agradecen los aportes de Harald Beyer y Carmen Le Foulon.

reabrir el debate acerca de la conveniencia tanto de eliminar las pruebas de conocimientos específicos como de exigir pruebas en asignaturas no afines a las carreras o áreas de estudio a las que se postula, e investigar la factibilidad de adaptar nuevas pruebas o secciones de pruebas extranjeras.

La construcción de instrumentos de evaluación requiere de procedimientos sistemáticos y rigurosos. En las pruebas con altas consecuencias, como es el caso de los exámenes de admisión a las universidades, estas exigencias cobran aun mayor relevancia. El proceso se inicia con la identificación del propósito de la evaluación y finaliza una vez que se demuestra que la prueba sirve para los objetivos para los cuales fue diseñada. Diferentes estándares internacionales sobre la construcción de pruebas definen una serie de pasos que deben respetarse al elaborar instrumentos de evaluación que implican consecuencias serias para los individuos. Respetarlos da garantías a los usuarios de que se someterán a un proceso justo y confiable. A continuación se describen brevemente dichos pasos y luego se analiza la posibilidad de cumplirlos que tienen las Pruebas de Selección Universitaria (PSU) 2003 en los plazos que el Consejo de Rectores le ha fijado.

I. ETAPAS DEL DESARROLLO DE PRUEBAS DE SELECCIÓN A LA UNIVERSIDAD¹

Las pruebas que se utilizan en educación pueden dividirse en dos tipos según las implicancias de sus resultados para el estudiante: de altas y bajas consecuencias. Los exámenes de admisión a la universidad se consideran de altas consecuencias, ya que a partir de sus resultados se toman decisiones trascendentales para las personas.

Con el objeto de regular el proceso de construcción de pruebas y garantizar un uso adecuado de los resultados, la comunidad de expertos en el tema ha elaborado estándares que prescriben lo que se considera una buena práctica en el ámbito del desarrollo y uso de instrumentos de evaluación. El documento que los recoge es el *Standards for Educational and Psychological Testing*, elaborado por la American Educational Research

¹ El listado de etapas que se presenta a continuación se basa en el documento elaborado por Carmen Le Foulon y Francisca Dussaillant, "Desarrollo de Pruebas Estandarizadas" (2002).

Association (AERA), la American Psychological Association (APA) y el National Council on Measurement in Education (NCME). El *Standards for Quality and Fairness*, desarrollado por el Educational Testing Service (ETS), se alinea con los anteriores y los refuerza. En adelante nos referiremos a ellos como estándares de la AERA y del ETS. Según el primer documento, “mientras más altas son las consecuencias de una prueba, más importante resulta que las inferencias que se hagan a partir de ella sean avaladas con evidencia sólida y de calidad técnica. En particular, cuando las consecuencias son individuales y serias, y cuando se toman decisiones importantes basadas en el desempeño en las pruebas, el instrumento necesita exhibir estándares elevados”².

Según los estándares, las etapas mínimas que se deben respetar al elaborar este tipo de pruebas son las siguientes:

1.1. Desarrollo de las especificaciones de la prueba

En la etapa inicial de la planificación de una prueba se debe definir la suma de sus características. Las especificaciones contienen el propósito de la prueba, la fundamentación que avala la selección de contenidos, destrezas y habilidades a ser medidas, la tabla de especificaciones de los contenidos a evaluar, la definición del formato de las pruebas y una anticipación de las consecuencias que se pueden esperar. “Idealmente las especificaciones deben ser tan completas que dos constructores de prueba operando independientemente sobre la base de ellas debieran producir instrumentos comparables e intercambiables, que difieran tan sólo en la muestra de preguntas utilizadas”³. A continuación se detallan cada uno de los aspectos que deben definirse en la etapa de especificación de la prueba.

Descripción clara de los propósitos de la prueba. Es de vital importancia identificar el o los propósitos principales para los cuales serán utilizados los resultados de la prueba. En este nivel de análisis se debe estudiar la factibilidad de abordar en una sola prueba los distintos objetivos que se están proponiendo. En ocasiones los objetivos pueden ser incompatibles, lo que hace necesario evaluar si es posible que todos ellos se logren a partir de una misma evaluación. En el caso de las pruebas de selección a la universidad que nos ocupan, pretender evaluar la calidad de la educación media y

² Véase Estándar 13, en AERA, APA, NCME, *Standards for Educational and Psychological Testing* (1999), p. 139.

³ S. Tinkelman, “Planning the Objective Test” (1971).

seleccionar a los alumnos a la universidad puede no ser viable. En un caso, el diseño apunta a la selección de alumnos para programas educacionales altamente competitivos, por lo tanto, debe focalizarse en la capacidad de adquirir conocimientos avanzados y en la evaluación de las habilidades para cursar estudios superiores; en el caso que se buscara evaluar el logro de los objetivos de la educación media, la prueba debiera cubrir el dominio de todo el rango de contenidos independientemente de la relación de éstos con los estudios universitarios.

Fundamentación del tipo de prueba. En la etapa de especificación de la prueba se debe aportar evidencia lógica, teórica y empírica que apoye las inferencias que se harán a partir de los puntajes obtenidos en ella. Si se trata de pruebas que evalúan conocimientos y destrezas, tiene que quedar claramente estipulado cuán representativa será la muestra de tareas que se exigirán del universo total de tareas que se quiere medir. Asimismo, en las pruebas que apuntan a habilidades o disposiciones debe fundamentarse el constructo teórico sobre el que se basa la medición, es decir, los supuestos que avalan la relación entre los indicadores indirectos escogidos y la habilidad misma. Cuando la evaluación pretende predecir desempeño futuro debe quedar fundamentada la evidencia lógica y empírica que sustenta la elección de los indicadores utilizados.

En esta etapa tiene que discutirse la delimitación de los constructos⁴ y dominios a evaluar; por ejemplo, hay que responder preguntas del tipo: ¿El dominio de la biología tiene alguna relación con el desempeño posterior en carreras del área de las ingenierías? ¿La matemática de IV año de educación media incluye derivadas? ¿Incluye aplicación de conceptos a la resolución de problemas o sólo reconocimiento de definiciones? ¿La habilidad verbal incluye comprensión lectora y manejo de vocabulario? En esta tarea, la consulta a expertos debe ser amplia para recoger información actualizada sobre cómo se definen y operacionalizan las habilidades a evaluar y sobre lo que consideran relevante de medir en cada una de las disciplinas. La calificación, relevancia de la experiencia y las características demográficas de los jueces deben quedar documentadas⁵.

⁴ El término constructo tiene una connotación amplia que se refiere a los conceptos o características que un test pretende medir. La connotación más estrecha reserva el término para características no observables directamente, pero que se pueden inferir a partir de un conjunto de observaciones interrelacionadas. Un ejemplo de constructo entendido en este sentido más estricto sería la motivación, ya que sólo se la puede observar a partir de un conjunto de disposiciones y acciones del individuo.

⁵ Véase Estándar 3.5 en AERA, APA, NCME, *Standards for Educational and Psychological Testing* (1999).

El marco teórico de la prueba debe servir de guía para las etapas subsecuentes de validación del instrumento. La definición precisa del universo de dominios a evaluar, la especificación clara de los constructos y la identificación de las variables de criterio que se utilizarán para contrastar las predicciones con la realidad, entregan los parámetros con los cuales los jueces expertos tendrán que revisar las pruebas finales.

Definición de los dominios a evaluar. En esta etapa se traduce operacionalmente el marco teórico en una definición detallada de los dominios, subdominios y tipos de destrezas a evaluar. Luego se explicitan en una tabla que contiene el peso que se le asignará a cada área. Ello resulta particularmente importante para el proceso de validación de contenidos y en los casos en que habrá un equipo relativamente grande de profesionales desarrollando ítems.

Formato de la prueba y de los ítems. Se debe definir el formato de la prueba (por ejemplo, papel o computador), el número de ítems a incluir por tarea y el número de ítems en cada área. También es importante especificar el nivel apropiado de lenguaje (extensión de las lecturas, tipo de lenguaje a utilizar) y se debe definir el formato de todas las preguntas (ensayo, opción múltiple, etc.) y las tareas a realizar (comprensión de lectura, completación de oraciones, eliminación de la incorrecta, etc.). Al nivel de los ítems es necesario establecer un límite de palabras para los enunciados, así como definir las características de las opciones y de los distractores (en el caso de preguntas de selección múltiple). Otro tema de importancia dice relación con las instrucciones de la prueba. Se debe definir de antemano los requerimientos para éstas, es decir, si bastará con algunas preguntas de ejemplo al inicio de cada sección de la prueba o si será necesario confeccionar pruebas completas de muestra para que la población esté al tanto del tipo de preguntas que deberá enfrentar.

Características psicométricas. Las especificaciones de la prueba deben incluir, entre otros, el nivel de dificultad de ésta, la distribución de dificultad de los ítems, directrices para evaluar la homogeneidad de los ítems⁶ y una descripción de los requerimientos para el *equating*⁷ (métodos, número de preguntas ancla, distribución de éstas, etc.). También se debe estimar el número de ítems a utilizar y el tiempo total de la prueba.

⁶ Es decir, verificar que los ítems en un determinado subgrupo son de la misma naturaleza.

⁷ Mecanismo que permite comparar los puntajes con los de otras formas paralelas de la misma prueba o con aplicaciones en distintos años.

Especificaciones de equidad. Antes de desarrollar la prueba, es importante definir las directrices para abordar las diferencias culturales y de género. Es importante definir cómo se asegurará la representatividad de los diferentes subgrupos al momento de discutir las preguntas y los análisis estadísticos que evalúen diferencias entre grupos (Differential Item Functioning, DIF, por ejemplo). También se debe discutir si la prueba supondrá la exposición a experiencias de aprendizaje similares o si justamente se quiere evaluar las diferencias de oportunidades.

El proceso de desarrollo de las especificaciones debe estar sujeto a continuas revisiones durante el proceso de construcción de las pruebas.

1.2. Construcción de los ítems

La construcción del conjunto inicial de ítems implica las siguientes actividades:

- a) Selección de un formato de ítems apropiado y verificación de que el formato sea adecuado para los examinados.
- b) Selección y entrenamiento de quienes redactarán los ítems.
- c) Producción de un conjunto grande de ítems.
- d) Revisión inicial de los ítems por expertos. En el caso de pruebas de alternativas, se deben revisar los siguientes aspectos:
 - Claridad y precisión: el enunciado debe definir claramente el problema o tarea, medir un solo concepto, contener suficiente información para resolver el problema sin que se requiera mirar las alternativas, ser consistente gramaticalmente con las alternativas y, en lo posible, no contener negaciones.
 - Relevancia y pertinencia de acuerdo a la tabla de especificaciones: las preguntas deben calzar con las habilidades y conocimientos que se busca evaluar.
 - Fallas técnicas en la construcción de los ítems: en este punto se revisa la calidad de los distractores, se analiza la plausibilidad de los mismos, es decir que representen en la medida de lo posible concepciones erróneas de la pregunta. En este sentido también es importante verificar que las alternativas no ofrezcan pistas que permitan responder las preguntas sin tener los conocimientos necesarios; por ejemplo, revisar que ninguna resalte por falta de concordancia gramatical.
 - Apariencia de sesgo: las preguntas no deben contener un lenguaje ofensivo, alarmante o que haga referencia a un subgrupo de los alumnos que rinden la prueba.

- e) Realizar pruebas *preliminares de los ítems* y revisarlos nuevamente a la luz de los parámetros anteriores. Esta tarea requiere probar los ítems en una muestra pequeña de individuos. La idea es hacer un *focus group* para detectar tempranamente posibles problemas de los ítems.

Los pasos para la construcción de los ítems son iterativos. Un ítem debe pasar múltiples revisiones antes de integrar la muestra de ítems que se prueban en la aplicación experimental.

- f) Realizar una *aplicación experimental* o piloto de los ítems en una muestra de la población de individuos que rendirán finalmente el test. La aplicación piloto involucra la administración de los ítems, ensamblados en formatos de prueba similar al definitivo, a una muestra representativa de los individuos que rendirán la prueba⁸. Esta aplicación busca determinar las propiedades estadísticas de los ítems y, cuando corresponda, eliminar los ítems que no cumplen con los criterios preestablecidos. En esta etapa se deben probar alrededor del triple de las preguntas que quedarán finalmente.

Se insiste en que la evaluación de los ítems debe realizarse en contextos lo más parecidos a los de las evaluaciones reales. La motivación de los alumnos, la preparación de los alumnos, el nivel de dificultad de la prueba (estimado a pulso en la evaluación de expertos y en las pruebas preliminares), la extensión de la prueba, el orden de las preguntas y el tiempo total deben ser lo más similares posible para que los índices estadísticos obtenidos no cambien con respecto a la prueba experimental. El equipo que desarrolla la prueba debe demostrar que las diferencias en las condiciones de administración de las pruebas experimentales con las finales, si es que las hay, no distorsionarán el comportamiento estadístico de los ítems⁹.

La aplicación piloto es una etapa importante dentro del proceso de construcción de ítems, pero no sustituye a los estudios de validez de las pruebas. Hambleton y otros (1991) señalan que aun cuando se utilicen sofisticadas técnicas estadísticas basadas en la Teoría de Respuesta al Ítem (IRT), ello no es garantía de que la prueba sea “técnicamente buena”. Los métodos basados en IRT no corrigen problemas tales como pruebas desalineadas de los objetivos propuestos o ítems de contenidos irrelevantes que se comportan bien estadísticamente.

⁸ En Teoría Clásica se exige una muestra representativa y en Teoría de Respuesta al Ítem (IRT) se exige una muestra grande y heterogénea de individuos.

⁹ Véase Estándar 3.9 en AERA, APA, NCME, *Standards for Educational and Psychological Testing* (1999).

1.3. Ensamblado de la prueba

De los formatos testeados en la prueba piloto, se eligen los ítems que quedarán en las versiones definitivas. Esta selección debe cumplir con la tabla de especificaciones y los parámetros psicométricos definidos en la etapa inicial de la prueba. Para ello, se utilizan los índices de dificultad obtenidos empíricamente en la prueba experimental descrita en el punto anterior. Los formatos finales no son esencialmente distintos de los utilizados en la prueba piloto en cuanto a su extensión y ordenación, pero esta vez el nivel de dificultad de la prueba responde a una estimación empírica que garantiza una versión que ordene a los individuos como se desea.

1.4. Investigar la confiabilidad y validez de la prueba

La validez se refiere al grado en que la evidencia empírica y la teoría respaldan la interpretación de los puntajes de una prueba. Según los estándares de la AERA, la validación es la consideración fundamental del proceso de desarrollo y evaluación de una prueba. Es la que garantiza que la prueba mida lo que dice medir. Aplicar una prueba sin estudios de validez es cuestionable, más aun tratándose de tests que implican altas consecuencias.

El proceso de validación implica la acumulación de evidencia sólida y científica para el modo en que se interpretarán los puntajes de la prueba, ya que no se validan los instrumentos sino que el uso que se hace de ellos. Cada intención de uso debe ser validada en su propio mérito¹⁰.

La validación se inicia en la etapa de especificación de la prueba cuando se delimitan explícitamente el propósito de la prueba y sus usos, así como las inferencias e interpretaciones que se harán de ella. Las decisiones acerca de la clase de evidencia que es más relevante para cada prueba deben quedar estipuladas en un conjunto de hipótesis que indiquen sucintamente los supuestos teóricos que se están asumiendo. Por ejemplo: “que los examinados que obtienen puntajes altos tendrán mayores probabilidades de éxito en cursos avanzados que los que obtienen puntajes bajos”, o “un buen puntaje en la prueba de matemática indica que el alumno domina todos los contenidos que son prerrequisito para un curso avanzado”. Estas definiciones son las que proveen el marco de referencia que guiará los estudios de validez.

¹⁰ Véase Estándar 1 en AERA, APA, NCME, *Standards for Educational and Psychological Testing* (1999).

La validación es un ejercicio conjunto de los equipos que elaboran las pruebas y de los que las utilizarán. Los que desarrollan las pruebas son responsables de la elaboración de argumentos sólidos que avalen los usos propuestos de la prueba y también les corresponde reunir la evidencia teórica y empírica. Los usuarios deben evaluar la evidencia que justifica el uso de la prueba en contextos particulares¹¹.

Los estándares de la AERA distinguen distintas fuentes de validez, cada una ilumina aspectos esenciales que deben ser estudiados cuando se quiere afirmar que una prueba mide lo que realmente dice medir. Actualmente se postula un concepto unitario de la validez donde la acumulación total de la evidencia sirve para construir un argumento sólido que respalde la prueba.

Según la AERA la evidencia debe provenir de los análisis siguientes¹²:

Análisis de los contenidos. Esta categoría dice relación con el grado en que los ítems incluidos en la prueba representan bien los contenidos que la prueba desea medir. En el caso de una prueba de admisión, se requiere documentar que los ítems incluidos representan cabalmente los contenidos descritos en la tabla de especificaciones. La evidencia de validez correspondiente a esta categoría se basa, esencialmente, en el juicio de expertos que se pronuncian acerca de la idoneidad de los ítems incluidos en la prueba.

Análisis de los procesos utilizados por los evaluados para responder la prueba. En la etapa de revisión preliminar de los ítems y en la etapa de revisión por jueces expertos se debe analizar la forma en que los alumnos responden los ítems. El análisis tiene que proveer evidencia de que la forma de enfrentar el problema planteado responde al constructo que se está midiendo. Por ejemplo, si se plantea que una sección de una prueba mide razonamiento matemático, es importante determinar si de hecho los alumnos están razonando o sólo están aplicando algoritmos estándares.

Análisis de la estructura interna de la prueba. Se debe demostrar el grado en que la totalidad de los ítems de una prueba y las secciones de la misma se relacionan con el o los constructos de la prueba. En los análisis preliminares se analiza la concordancia de cada ítem con el constructo de la prueba; en esta nueva instancia se vela por la coherencia global del instrumento. Por ejemplo, si se postula que una prueba mide una sola dimensión, se debe demostrar su homogeneidad.

¹¹ *Ibíd.*

¹² Para una descripción detallada de estas líneas de validez véase Estándar 1, en *ibíd.*, pp. 11-17.

Análisis de la relación de los puntajes de la prueba con variables externas a la prueba. Entre las variables externas se incluyen los estudios de validez predictiva, que relacionan los puntajes de la prueba con la medición de la variable de criterio que se pretende predecir. Los estudios de validez concurrente también se consideran en esta categoría; éstos están orientados a encontrar una relación directa con otros instrumentos que miden constructos similares. En este sentido, también aporta evidencia la comparación de los resultados con pruebas que miden constructos relativamente distintos. En el caso de esta validación discriminante, la relación entre los puntajes de una y otra prueba debieran ser bajos.

Análisis de las posibilidades de generalizar la utilidad de la prueba a otros contextos. Cuando la prueba se utiliza en contextos distintos a los originales se debe demostrar que éstos no afectan a la validez del instrumento. Por ejemplo, si la naturaleza de las habilidades y conocimientos requeridos para cursar con éxito el primer año de universidad varían sustancialmente, es probable que la validez predictiva de la prueba se modifique. En este caso se deben hacer los estudios correspondientes para ajustar la prueba a los nuevos requerimientos si es que los indicadores de predictibilidad bajan significativamente.

Análisis de las consecuencias producidas por las pruebas. Los estándares de la AERA exigen demostrar que el uso de las pruebas logra los beneficios que sus constructores anuncian. De igual modo deben reunir evidencia de que no producen efectos negativos. Por ejemplo, si se asevera que una prueba de admisión a la universidad mejorará los índices de logro académico en la educación media, se deben conducir los estudios que lo demuestran¹³.

En el caso de las pruebas de selección a la universidad, Shepard (1993)¹⁴ afirma que se debe partir por identificar el propósito explícito de las pruebas que en el caso de las de admisión es seleccionar estudiantes que tengan posibilidades de éxito en la universidad. Esta autora plantea que, dado lo anterior, lo primordial y más importante tratándose de un argumento de validación, es que la prueba debe demostrar su poder predictivo, es decir, que se observa correlación entre el puntaje obtenido y el rendimiento universitario. Pero tal relación no debe ser la única dimensión a explorar: es necesario demostrar que el contenido de las pruebas evalúa las habilidades

¹³ Véase AERA, APA, NCME, *Standards for Educational and Psychological Testing* (1999), p. 4.

¹⁴ L. Shepard, "Evaluating Test Validity" (1993), p. 19.

que son prerequisite para lograr éxito en la universidad. Asimismo, hay que proveer evidencia de que las pruebas no presentan sesgos que perjudiquen a algún grupo particular debido a la forma en que se pregunta. Si hay diferencias entre los individuos, éstas no pueden ser atribuibles a sesgos sino a varianza verdadera en el atributo medido.

Los procesos de validación toman tiempo. Por ejemplo, en Estados Unidos en el año 2002 se tomó la decisión de modificar el SAT I, una de las pruebas de admisión a la universidad que allí se aplican. Aunque se trata de cambios menores¹⁵, la nueva versión entrará en vigencia el año 2005, es decir se darán un plazo prudente para realizar los estudios de validez necesarios. Otro tanto ha ocurrido con la introducción de la prueba SAT en Singapur. En Suecia los estudios experimentales de la SweSAT comenzaron en 1973 y ella se aplicó por primera vez en 1977. Los cambios marginales introducidos desde ese entonces se han tardado, en promedio, tres años. Los cambios han sido sugeridos por un Consejo Internacional. Desde 1996 se permite en ese país testear los nuevos ítems y eventuales nuevas secciones en la misma prueba de selección.

Si bien se puede reunir evidencia de validez a partir de las primeras etapas de la construcción de un instrumento, el verdadero esfuerzo en este sentido no se puede desarrollar sino hasta que el instrumento está en su forma final. La validez predictiva requiere, además, que el instrumento se aplique y opere por un tiempo para que se recojan los indicadores de la variable de criterio. Esto último introduce un dilema, ya que no se debiera aplicar instrumentos que acarreen consecuencias serias para los alumnos sin haber verificado previamente su validez, pero a la vez esta evidencia no se puede obtener sin aplicar esos instrumentos. Una de las soluciones a las cuales se recurre es la de implementar los cambios gradualmente. Los nuevos instrumentos pueden pilotarse en paralelo a los antiguos, sin aplicar las consecuencias que normalmente se asociarían a la prueba. Éste es el procedimiento que se empleó al cambiar el antiguo sistema de admisión chileno, vía Bachillerato, por el de la Prueba de Aptitud. El otro sistema es el que ha empleado el College Board en el proceso de modificación del SAT I. Allí se busca introducir secciones y contenidos nuevos a los instrumentos antiguos. Sólo en el momento que se comprueba que dichas secciones tienen igual o mayor poder predictivo que las anteriores se las reemplaza definitivamente. En este ejercicio de cambios graduales, en el que se cambian partes de la prueba y no el sistema total, se corren menos riesgos

¹⁵ En la prueba de lenguaje se elimina la sección de analogías y se incorporan ítems de redacción previamente aplicados y probados por años en la prueba específica de redacción (SAT II Writing). En la de matemática se mantiene el acento en razonamiento y se agregan contenidos de álgebra II.

de cometer injusticias con los alumnos que se someten al proceso de selección.

Para finalizar el tema de validez, los estándares del ETS insisten en que los estudios de validación deben repetirse a lo largo de la historia de aplicación de las pruebas. Sugieren que no pasen más de cinco años entre dichos análisis.

Otro aspecto relevante a estudiar es la confiabilidad de los instrumentos. Ésta se refiere a la consistencia de las mediciones cuando ellas se repiten en la misma población o grupos. Se supone que los rasgos que se miden son relativamente estables en el tiempo, a menos que exista una intervención directa para modificarlos. Si no se ha dado esa intervención, se estima que el comportamiento entre una aplicación y otra no debe variar sustancialmente. Una de las técnicas utilizadas para comprobar la confiabilidad consiste en dividir la prueba, una vez aplicada, en dos partes equivalentes y analizar las variaciones de desempeño de la misma población en cada mitad de la prueba. En general, no se acepta más de un 10% de variación en el desempeño de ambas partes. Los estudios de confiabilidad son una consideración necesaria para legitimar una prueba, pero no reemplazan los estudios de validez antes descritos.

1.5. Garantizar la seguridad de la prueba

En las pruebas de altas consecuencias debe quedar especificado cómo se garantizará que los resultados sean confiables y no producto del acceso fraudulento a las preguntas. En este sentido, cobra importancia el cuidado con las filtraciones de preguntas. Se deben tomar los resguardos necesarios para que no se sustraigan pruebas o ítems.

En pruebas que se repiten de año en año es más fácil que las preguntas se filtren. Los interesados en darlas en el futuro pueden recoger antecedentes de las preguntas con los que las rindieron en años anteriores. Para evitarlo se hacen formas paralelas que permitan la comparación entre pruebas sin repetir las preguntas. Sin embargo, para lograr la equivalencia entre las pruebas se debe conservar un número de ítems ancla que se aplican en ambas mediciones. Estas preguntas representan una proporción menor de la prueba pero quedan vulnerables a la filtración. Otro tanto sucede con las preguntas que se testean experimentalmente en el contexto de la medición del año anterior a la prueba¹⁶. El mejor antídoto para este problema es crear

¹⁶ Los alumnos que dan la prueba deben contestar preguntas que corresponderán a los ítems de mediciones posteriores. Los alumnos no reciben puntaje por estas preguntas pero ellos no lo saben. Esta política se aplica porque permite evaluar los ítems en un contexto de motivación real. El problema con este procedimiento es que los alumnos pueden ponerse de acuerdo para anotar las preguntas y entregárselas a la futura generación que será evaluada.

grandes bancos de ítems que disminuyan el incentivo de conocer preguntas específicas, dada la baja probabilidad de que la pregunta filtrada corresponda a una de las preguntas que aparecerá en la prueba.

1.6. Presentación y validación de la prueba ante el público

La etapa final de la construcción de una prueba es el estudio de cómo se explicará al público la lógica que sustenta la prueba, la interpretación de los resultados, y la evidencia de validez y equidad. La disposición a aceptar los veredictos de las pruebas se relaciona con el grado de información que tienen los usuarios acerca de la naturaleza de la prueba y de su comprensión de aquello que las sustenta. En este sentido, la información técnica debe ser transparente.

En el caso de las pruebas de admisión, un factor importante es dar a conocer a tiempo los facsímiles de la prueba. En parte, porque ayudan a la comprensión de la prueba, pero también porque los alumnos deben familiarizarse con el formato de las secciones del examen. Conocer el tipo de ítems es relevante porque los resultados pueden distorsionarse si los alumnos no están familiarizados con ellos. En ese caso, la prueba estaría evaluando la capacidad para enfrentar tareas nuevas más que las destrezas que se pretende medir, lo cual confundiría la interpretación de resultados. En este sentido, se debe cuidar que los alumnos tengan igualdad de oportunidades para practicar y familiarizarse con el formato de las pruebas. Según los estándares de la AERA, “Los examinados tienen derecho al acceso igualitario a los materiales elaborados por los patrocinadores de las pruebas que describen el contenido y propósito de la evaluación, así como al material diseñado para familiarizar y preparar a los alumnos para la prueba”¹⁷.

1.7. Documentación del proceso

Los estándares revisados proponen que cada una de las etapas de elaboración de las pruebas quede debidamente documentada. Se considera fundamental recopilar los antecedentes de cada uno de los pasos de las etapas descritas anteriormente para facilitar la revisión y validación de las mismas, así como para proveer información a los usuarios que les facilite la interpretación de los juicios que se desprenden de los resultados de las evaluaciones¹⁸. El listado de jueces externos que participan en el proceso

¹⁷ Véase Estándar 7 en AERA, APA, NCME, *Standards for Educational and Psychological Testing* (1999), p. 75.

¹⁸ Véase Estándar 6 en *ibídem*.

de validación de constructo y de contenido debe quedar debidamente registrado, así como los informes que ellos emiten. “Cuando el proceso de validación descansa en parte en la opinión de jueces expertos..., los procedimientos para seleccionar dichos expertos deben ser plenamente descritos. La calificación y experiencia deben acreditarse”¹⁹.

Finalmente, los estándares de la AERA consideran que las exigencias establecidas para la confección de pruebas deben ser cumplidas antes del uso operacional de las pruebas²⁰.

II. ANÁLISIS DE LA PSU 2003 A LA LUZ DE LAS ETAPAS DE DESARROLLO DE UNA PRUEBA

En Chile, en el marco de los proyectos concursables FONDEF, se elaboró un nuevo sistema de pruebas de admisión a la universidad, el que supuestamente se comenzaría a aplicar a fines de 2002. Éste nuevo sistema, conocido como SIES (Sistema de Ingreso a la Educación Superior), reemplazaría a la Prueba de Aptitud Académica y a las pruebas de Conocimientos Específicos vigentes hasta entonces. Sin embargo, tras ser objeto de diversas críticas durante el año 2002, el SIES fue descartado y en su reemplazo se ha dispuesto desarrollar, y aplicar a fines de 2003, una nueva batería de pruebas denominada PSU (Pruebas de Selección Universitaria).

El objetivo de esta sección es analizar la PSU 2003 a la luz de las etapas de desarrollo que deben cumplir las pruebas de selección para la educación superior. Ahora bien, el análisis deberá limitarse a los pocos antecedentes disponibles, ya que no hay documentos públicos que expliciten la configuración de la PSU 2003. En efecto, la Comisión Técnica encargada de elaborar la PSU 2003 está entregando paulatinamente los lineamientos que se adoptarán, pero no se conocen públicamente otros documentos que contengan antecedentes más completos. Es cierto que el proyecto SIES, que es uno de los insumos de esta nueva prueba de admisión, entrega una fundamentación, pero esa fundamentación es adecuada en el contexto de un informe para concursar a fondos de investigación y no cumple con los requisitos de especificación de la etapa inicial de una prueba de estas características.

Conforme a lo anterior, primero se realizará un recorrido por las pruebas que se utilizarán como insumos para la PSU (esto es, la PAA, las pruebas de Conocimientos Específicos y el SIES), y luego se procederá a examinar la PSU 2003.

¹⁹ Véase Estándar 1.7 en *ibídem*.

²⁰ *Ibídem*.

2.1. Prueba de Aptitud Académica (PAA)²¹

Etapas de especificación de la prueba	La Prueba de Aptitud Académica (PAA) se basó en sus inicios en el examen de admisión a las universidades norteamericanas SAT I, por lo tanto contaba con un marco definido de especificaciones. Las Pruebas de Conocimientos Específicos, más vinculadas a contenido, también siguieron el modelo de las pruebas SAT II con adaptaciones al currículo nacional.
<i>Propósito de la prueba</i>	El propósito de la PAA está claramente definido. Su objetivo central es la identificación de los alumnos que tienen mayores posibilidades de rendir con éxito los estudios universitarios. Se expresa en la probabilidad de egresar oportunamente y de obtener mejores notas en la universidad.
<i>Fundamentación</i>	Al momento de instaurar la PAA se contaba con los estudios de validación empírica realizados en Estados Unidos, los cuales avalaban a esta prueba como un buen predictor del desempeño académico de los universitarios. El constructo de aptitud académica tenía un vasto apoyo teórico y buenas definiciones operacionales de lo que se quería medir. Todo lo cual facilitaba la evaluación cualitativa del contenido de la prueba de parte de los jueces expertos. Con el tiempo, el SAT I ha acumulado evidencia de su poder predictivo y se ha modificado ligeramente la definición del constructo que subyace en la prueba. Cambió el concepto de aptitud por el de razonamiento para hacerlo más comprensible al público y quitarle la connotación de heredabilidad inmutable. El nuevo concepto de razonamiento busca sintonizar con la idea de que es una destreza que se puede adquirir lenta y transversalmente en toda la formación escolar. En Chile, la PAA se validó empíricamente antes de aplicarse. Durante cuatro años de marcha blanca se estudió su capacidad predictiva. Hasta el año 1992, los estudios de validez predictiva se

²¹ Para más detalles, véase G. Donoso, M. A. Bocchieri, E. Ávila y otros, *El Sistema de Admisión: Orígenes y Evolución. Resultados del Proceso de Admisión 1999* (1999).

<i>Definición de los dominios a evaluar</i>	hacían periódicamente. No se continuó con esta política en los años siguientes, transgrediéndose así los estándares de validación que sugieren realizar este tipo de análisis al menos cada cinco años. El último estudio formal de validez de constructo data del año 1987; sin embargo, posteriormente se hicieron análisis internos para modificar secciones de la prueba. Los últimos cambios en la definición del constructo realizados por el SAT I no han sido abordados aún por la PAA. La prueba de aptitud define claramente los dominios y subdominios a evaluar. Están estipulados para cada una de las pruebas y las secciones responden a ellos.
<i>Formato de la prueba</i>	La prueba ha seguido un patrón común en sus evaluaciones, por lo tanto hay modelos previos para la construcción de nuevas pruebas. Periódicamente se realizan estudios para reemplazar secciones y tipos de preguntas. Éstos se comunican a través de un boletín que llega anualmente a los colegios y aparecen debidamente ejemplificados en los facsímiles que se entregan a cada postulante.
<i>Características psicométricas</i>	La capacidad discriminativa de los ítems de la prueba se realiza en cada aplicación con pruebas experimentales y aplicando teoría clásica de medición. A la Prueba de Aptitud Académica se le exige que discrimine en cada uno de los niveles de habilidad, buscando una distribución de los resultados cercana a una curva normal.
<i>Especificaciones de equidad</i>	En la PAA se conduce un análisis del comportamiento de los ítems según dependencia administrativa²² de los establecimientos. Las preguntas que presentan las diferencias menores de desempeño entre dependencias, en dificultad, discriminación y porcentaje de omisión son las que se integran al banco de ítems que serán empleados en

²² La dependencia administrativa se refiere al tipo de sostenedor y sistema de financiamiento de cada establecimiento. Según la dependencia administrativa, se consideran tres tipos de establecimientos: particulares pagados, particulares subvencionados y municipalizados.

	las pruebas finales. De esta manera se introduce un grado de control del sesgo cultural que pudiera contener la prueba.
Construcción de los ítems	La construcción de los ítems de la PAA se ve facilitada por la larga tradición de la prueba. Desde sus inicios se contó con un modelo en el cual apoyarse y en la actualidad el banco de preguntas es considerable. Las comisiones se coordinan anualmente por disciplinas y cuentan con un conjunto de revisores que revisan los ítems en cada una de las etapas. Los ítems se analizan de acuerdo a la Teoría Clásica de medición, y no se aplica IRT.
Estudios de confiabilidad y validez <i>Validez predictiva</i>	El último estudio de validez de constructo, publicado por el Departamento de Evaluación, Medición y Registro Educativo (DEMRE), de la PAA se hizo el año 1987. En 1989 se realizan estos análisis para las Pruebas de Conocimientos Específicos. En ellos se concluye que la pruebas son confiables y miden lo que pretenden medir. Los estándares recomiendan que estos estudios se repitan cada cinco años, requerimiento que no se ha cumplido. Los estudios de validez predictiva se han realizado a lo largo de la aplicación de la prueba y confirman la capacidad de predecir éxito académico. El último análisis de predictibilidad publicado por el DEMRE corresponde al proceso de admisión 1992. A raíz de los cuestionamientos a la prueba, se realizaron estudios de predictibilidad en distintas universidades durante el año 2002²³. Los resultados muestran que los indicadores obtenidos son similares a los encontrados en EE.UU. para la prueba SAT I, los cuales son aceptados por los expertos como un indicador de una buena capacidad predictiva²⁴.

²³ Véase R. Fischer y A. Repetto, “Método de Selección y Resultados Académicos (2002). También B. Vial y R. Soto, “¿Predice la PAA el Rendimiento o el Éxito en la Universidad?” (2002).

²⁴ Los indicadores de predictibilidad de las pruebas SAT I y SAT II se pueden encontrar en W. J. Camara y G. Echternacht, “The SAT I and High School Grades: Utility in Predicting Success in College” (2000), y en L. Ramist, C. Lewis y L. McCamley-Jenkins, “Using Achievement Test/Sat II: Subject Tests to Demonstrate Achievement and Predict College Grades: Sex, Language, Ethnic, and Parental Education Groups” (2001).

<i>Estudios de sesgo</i>	La PAA corrige los ítems por apariencia de sesgo y por comportamiento errático del ítem durante la prueba experimental. Sin embargo no hace un análisis de comportamiento diferencial de los ítems como lo hace el SAT I.
<i>Estudio de las consecuencias de las pruebas</i>	Una de las críticas que se le han hecho a la PAA gira en torno a las señales contraproducentes que envía hacia la educación media. Se dice que los alumnos descuidarían el estudio de las materias del currículum que no están incluidas en la prueba. No se han realizado estudios empíricos al respecto. Los estándares recomiendan su realización.
Consideraciones de seguridad	La PAA no es equiparable de un año a otro, por lo tanto, por esa vía no se pueden filtrar las preguntas porque éstas no se repiten entre aplicaciones. A lo largo de su existencia, se han filtrado preguntas por otros medios, pero se han ido creando los mecanismos para evitar este tipo de episodios.
Validación de la prueba ante el público	La Prueba de Aptitud Académica ha contado con legitimidad entre los estudiantes. El proceso de admisión se ha reconocido como transparente y confiable.

2.2. Sistema de Ingreso a la Educación Superior (SIES)

Etapas de especificación de la prueba	Se conocen dos documentos públicos sobre el SIES, el “Informe de la Comisión Nuevo Currículum de la Enseñanza Media y Pruebas del Sistema de Admisión a la Educación Superior” ²⁵ , elaborado por la Comisión del mismo nombre convocada por el Ministerio de Educación, y el proyecto FONDEF, “Reformulación de las Pruebas de Selección a la Educación Superior” ²⁶ , dirigido por David Bravo y Jorge Manzi. Ninguno de
--	--

²⁵ Véase Comisión Nuevo Currículum de la Enseñanza Media y Pruebas del Sistema de Admisión a la Educación Superior, “Informe de la Comisión Nuevo Currículum de la Enseñanza Media y Pruebas del Sistema de Admisión a la Educación Superior” (2000).

²⁶ Véase David Bravo y Jorge Manzi, “Reformulación de las Pruebas de Selección a la Educación Superior” (proyecto FONDEF) (2000).

Propósito de la prueba	los dos documentos cumple los criterios de especificación que establecen los estándares para la elaboración de pruebas de altas consecuencias. El Informe analiza la problemática de las pruebas de admisión a la universidad y delinea un marco general de operación, pero éste dista de ser un marco conceptual para la formulación de una prueba porque carece del nivel de especificación necesario. Tampoco se presenta la evidencia teórica y empírica que avale las opciones escogidas. Por su parte, el proyecto FONDEF no compensa dichas falencias. En él se asevera que el diseño de las pruebas se basa en el marco orientador del Informe de la Comisión <i>ad hoc</i> convocada por el Ministerio de Educación y no se ahonda más en el tema. En el cronograma presentado en dicho informe no se le asigna tiempo a la fase preparatoria de definición del marco de especificaciones de la prueba. Aparentemente, dicha etapa se dio por finalizada con el trabajo de la Comisión. En el Informe de la Comisión se establecen una serie de objetivos que deben ser cumplidos por las nuevas pruebas. En primer lugar se busca “robustecer un sistema de selección efectivo, confiable y de alta legitimidad como el vigente, y contribuir, a través de las pruebas de tal sistema y las presiones y orientaciones que de hecho ellas establecen sobre los últimos dos años del nivel secundario, al logro de los nuevos objetivos curriculares de la educación media” (p. 4). También se busca “la producción de información estratégica para actores de diversos ámbitos, sobre el cumplimiento de los objetivos de aprendizaje tanto del sistema escolar como de la educación superior” (p. 7). En la sección de proposiciones de mejoramientos y cambios se insiste en el tema y se asevera que es necesario hacer un esfuerzo integral de adaptación de las pruebas para hacer “converger la presión de las evaluaciones del sistema de educación terciaria, con el currículum de la educación media, en forma más clara, directa y robusta que en el presente, para mejora de los resultados de la enseñanza media (EM) y de la preparación de los estudiantes de la EM”; también se plantean como necesarios “la evaluación de la educación media y los saberes y
-------------------------------	--

	habilidades de los egresados” (p. 42). Se afirma que las pruebas no pueden ser consideradas como un problema sólo de la educación superior, ya que “tienen una doble función: de selección y de evaluación de los resultados formativos de la educación media” (p. 42). En este sentido, el Informe afirma que “las nuevas pruebas deben referirse a los objetivos y contenidos del currículum de la educación media, en cada una de las asignaturas que consideren. El rediseño planteado debe considerarse como criterio orientador el que se trata de pruebas referidas al currículum oficial, y que, por tanto, las restricciones impuestas en los contenidos y habilidades medidas, por necesidades de discriminación de los instrumentos y bajos rendimientos actuales, deben ser levantadas” (p. 45). El Informe también plantea que de acuerdo a lo recogido por la Comisión respecto de la discusión e investigaciones internacionales sobre la materia, el giro planteado en la prueba podría lograr “consecuencias positivas sobre la equidad, al hacer disminuir el peso de las variables asociadas a lo que se describió como <i>coeficiente de herencia</i>, y acrecentar, en vez, las asociadas a la experiencia y el trabajo escolar”. De la lectura del Informe se desprenden, entonces, al menos cuatro objetivos: mejorar la eficiencia en la selección de postulantes a la universidad, mejorar la equidad del ingreso a la universidad, aumentar el rendimiento de los alumnos en enseñanza media, evaluar la educación media y entregar información acerca de los logros de este ciclo. Ni el Informe ni el proyecto FONDEF se detienen a discutir cómo se podría llegar a conciliar el logro de todos estos objetivos en una sola prueba. Técnicamente no es claro que las pruebas puedan evaluar la educación media y a la vez seleccionar bien a los alumnos para la universidad. Además, es difícil pensar cómo se puedan incluir en la prueba todos los temas relevantes del currículum independientemente del nivel de logro de los alumnos. Si hay tópicos del currículum que una amplia proporción de los alumnos no dominan, las preguntas que consideran esos temas no sirven para ordenar a los alumnos, ya que nadie las puede contestar. Al no tener capacidad discriminante son inútiles al momento de evaluar.
--	--

<i>Fundamentación</i>	En el Informe de la Comisión se critica brevemente la noción de aptitud sobre la que se basaba la PAA. Básicamente, se cuestiona el supuesto de que las aptitudes académicas se distribuyen normalmente, es decir, que serían relativamente independientes de variables tales como sexo, edad, nivel socioeconómico y cultural. También se cuestiona la estabilidad de las aptitudes en el tiempo, la cual se refiere a la idea de que las habilidades no son modificables con el entrenamiento y la madurez de los individuos. Junto a estas críticas se asevera que el avance de la psicología cognitiva demuestra que es prácticamente imposible medir el logro de las capacidades cognitivas de los alumnos sin contemplar en tal medición los contenidos sobre los que se aplican las capacidades cognitivas. Al final de un somero análisis, realizado en dos páginas, que cita dos estudios, se concluye que “lo más razonable es abandonar la conceptualización de la aptitud y basar, en cambio, el sistema de admisión a la enseñanza superior en una evaluación que incluya la combinación de procesos cognitivos y contenidos curriculares que forman parte de la experiencia regular de los estudiantes de enseñanza media” (p. 11). En otra sección del Informe se avala con la experiencia comparada la noción de que es válido seleccionar a los alumnos para la universidad basándose en la evaluación del logro del currículum de educación media. Se mencionan los casos de Inglaterra, Francia, Alemania, Suecia, Japón e Israel citando el libro de Britton y Raitzen de 1996. Sin embargo, no se hace un análisis de la evidencia entregada. Por ejemplo, no se toma en cuenta que en la mayoría de estos países la selección de alumnos por habilidad académica ha ocurrido con anterioridad durante su educación escolar al separar a los alumnos por líneas vocacionales. Tampoco se actualizaron los datos, por lo que no se recoge la evidencia de que Suecia e Israel han incorporado pruebas tipo SAT I en la selección de alumnos. También se olvida mencionar que las pruebas de currículum de estos países son de ensayo, y por lo tanto, incluyen las habilidades de razonamiento a través de la fundamentación y organización lógica de la exposición.
------------------------------	---

Con respecto al tema de equidad se afirma que la evidencia indicaría que las pruebas referidas al currículum son más susceptibles de ser afectadas por la experiencia escolar y que por tanto serían más equitativas. Sin embargo, no se hace referencia a que esta línea de evidencia indicaría a su vez que cuando la calidad de la educación es muy disímil los resultados de las pruebas referidas a currículum resultan más inequitativas que las referidas a razonamiento²⁷. En todo caso el Informe no entrega evidencia empírica en uno u otro sentido.

La evidencia y la discusión entregada en el proyecto FONDEF son escasas, por lo que se pueden considerar como un análisis preliminar del problema pero en ningún caso como un marco teórico que justifica y explicita el tipo de prueba que se elaborará. Falta una discusión más fundamentada acerca de las posibilidades que tiene una prueba referida al currículum nacional de ser un buen predictor de éxito en la educación superior; falta un análisis acerca de cuáles son los contenidos del currículum que tienen relación con las destrezas y conocimientos requeridos en la educación universitaria. Habría que contestar si todos los contenidos de la educación media están orientados hacia la universidad o si cumplen otras funciones, dado que la educación media es un fin en sí y no está concebida solamente como una etapa preparatoria para la universidad. Sería necesario demostrar que las pruebas hasta ahora aplicadas tienen menor valor predictivo que las que se pretenden implementar o que tienen igual valor predictivo pero menores consecuencias negativas secundarias. Habría que hacerse cargo de la evidencia de que cada vez más países están incorporando en sus baterías de selección, pruebas que tienen un fuerte énfasis en la evaluación de la capacidad de razonamiento, dado que los currículos han enfatizado el desarrollo de este tipo de destrezas.

En mayo de 2002, Rosas, Flotts y Saragoni publicaron el artículo “Modelo de Representación del

²⁷ Véase H. Beyer, “Las Nuevas Pruebas de Ingreso a la Universidad” (2002), pp. 17-20.

<i>Definición de los dominios a evaluar</i>	Conocimiento para las Nuevas Pruebas de Selección a las Universidades Chilenas”. En él, los autores formulan especialmente un modelo del funcionamiento cognitivo para guiar la construcción de las preguntas del SIES. Este marco conceptual, original para estas pruebas, no se puede considerar como un referente validado, ya que la investigación en el campo de la cognición está en pleno desarrollo y discusión. No es un modelo que cuente con el apoyo de la experiencia comparada y posiblemente tampoco ha sido discutido por la comunidad de expertos en el tema. Para considerarlo como el elemento eje en la confección de preguntas requeriría de un proceso de validación. La definición de los dominios a evaluar se hace en forma muy general, por lo que no satisface la exigencia de los estándares de detallar con precisión los contenidos y habilidades que se evaluarán. El Informe concluye que “Las nuevas pruebas procurarán medir competencias en áreas del currículum. El concepto de competencias alude a capacidades de desempeño en contextos simbólicos o prácticos determinados. El propósito de medición de los nuevos instrumentos no debiera centrarse exclusivamente en contenidos, ni tampoco en procesos o capacidades intelectuales de los postulantes. Las nuevas pruebas debieran orientarse a medir el desempeño de los estudiantes en situaciones problemáticas nuevas, empleando la combinación de los procesos cognitivos y contenidos que han desarrollado durante su experiencia regular en la enseñanza media, asociada directamente a su trabajo en las disciplinas del currículum”. (p. 44). Esta definición no basta para elaborar las tablas de contenidos y destrezas a evaluar y menos el peso que se le asignará a cada una de éstas. En primer lugar, porque el marco curricular chileno de la educación media, en las asignaturas de lenguaje, matemática, ciencias naturales y sociales no permite una interpretación unívoca de cuáles son los conocimientos y procesos cognitivos que los alumnos deben lograr en cada uno de los dominios. Éste no explicita bien la cobertura y profundidad de cada uno de los contenidos y destreza a desarrollar.
---	---

	Probablemente, las definiciones entregadas en el Informe no darían pie para que dos equipos elaboradores de pruebas, trabajando separadamente, construyeran pruebas similares. Lo cual demostraría que la etapa de especificación de los dominios de la prueba no se habría implementado bien. Cuando este paso no se cumple satisfactoriamente, los estudios de validez de contenido se dificultan porque no quedan estipulados claramente los criterios con los cuales se tienen que refrendar las pruebas.
<i>Formato de la prueba</i>	La definición del formato de la prueba no aparece ni en el Informe de la Comisión ni en el proyecto FONDEF. Ciertamente esto no correspondía definirlo en el Informe ni en el proyecto; sin embargo, en documentos posteriores debieron haber quedado estipulados. Las decisiones acerca del número de preguntas por evaluación, el tipo de secciones al interior de cada una de las pruebas, no se encuentran analizadas en dichos documentos.
<i>Características psicométricas</i>	En el proyecto FONDEF se describen los criterios psicométricos fundamentales para la selección de los ítems, se especifican las curvas de información que se buscará en la confección de las pruebas y la capacidad discriminativa que se exigirá a los ítems. También se explicitan los análisis de sesgo que se realizarán a las preguntas. Sin embargo, el proyecto no fundamenta el tipo de decisiones que toma. No queda claro por qué las curvas de información serán tan amplias (de hecho cubrirían casi todo el espectro posible de desempeño) y cómo se cuidará que el error de medición no se eleve considerablemente por este hecho. Probablemente esta especificación quedó fijada así por la doble función que se le asignó a la prueba de seleccionar a los alumnos a la universidad y de evaluar el logro en educación media. El marco de especificación de la prueba debiera analizar más claramente este punto y dar garantías de que los alumnos quedarán confiablemente jerarquizados de acuerdo a su desempeño y dentro de rangos de error aceptables.

<i>Especificaciones de equidad</i>	El Informe especifica que se realizará un análisis del sesgo potencial de las preguntas debido a variables tales como tipo de dependencia, modalidad de enseñanza, sexo y región.
Construcción de los ítems	A la etapa de construcción de las preguntas se le asigna un período de tres meses en el proyecto FONDEF. Este plazo es insuficiente si se considera que esta etapa requiere conformar los equipos de redacción; entrenar a los redactores de las preguntas; crear un <i>pool</i> amplio de preguntas de las pruebas y de las maquetas en 6 asignaturas; evaluarlas por expertos en cuanto a su claridad, relevancia y pertinencia respecto a las tablas de especificaciones, fallas técnicas, gramática, apariencia de sesgo, estimación de grados de dificultad; y reescribir las preguntas necesarias y volver a revisarlas. En el proyecto FONDEF no se estipula tiempo para la etapa de entrenamiento de los redactores de las preguntas y esto representa una deficiencia importante, ya que no basta el grado de conocimiento de la disciplina, sino que los redactores deben estar compenetrados con los objetivos generales de la prueba y deben tener conocimientos de la redacción de preguntas de opción múltiple. Esto es más preocupante, si se considera que no sólo cambian los parámetros de la prueba sino que también los equipos que elaboran la prueba. De partida se integra un número mayor de profesores de educación media y disminuye el número de profesores universitarios especialistas en cada asignatura. La conformación de los equipos de expertos que revisan los ítems debiera ser transparente y sus informes quedar debidamente documentados. El proyecto FONDEF asigna alrededor diez días para esta subetapa, tiempo claramente insuficiente para realizar un análisis riguroso del conjunto de las preguntas destinadas a la prueba y a las maquetas de difusión de la prueba.
Estudios de confiabilidad y validez <i>Confiabilidad</i>	En el proyecto FONDEF se afirma que se realizarán estudios de confiabilidad, pero no se detalla qué tipo de técnicas utilizarán.

<i>Validez</i>	El proyecto FONDEF no contempló en su cronograma la fase de validación de contenido de la prueba: no se consideraron plazos ni recursos. No se menciona la conformación de un equipo de expertos que evalúen las pruebas ni en la etapa experimental ni después de la primera aplicación formal de la prueba. Los estándares exigen que el criterio con el cual se elige a los jueces expertos quede claramente especificado y que se describa el proceso por el cual se llega a las conclusiones finales²⁸. El acopio de evidencia de validez de contenido debe darse en varias etapas de la construcción de la prueba, no basta con analizar los ítems independientes al momento de su elaboración sino que es necesario estudiar los prototipos de las pruebas completas para certificar que cumplen con la tabla de especificaciones generales, que están alineados con el propósito de la prueba y con la fundamentación teórica.
<i>Validez predictiva</i>	El proyecto FONDEF no menciona la posibilidad de realizar estudios de predictibilidad. Esta es una carencia severa ya que el objetivo central de una prueba de selección para la universidad es precisamente predecir el éxito en ella. Según los estándares de la AERA, si la prueba se utiliza para aseverar que los candidatos que tienen mayores puntajes se desempeñarán mejor en la universidad, los encargados de elaborar la prueba y los que la utilizan deben proveer información acerca de la asociación que existe entre los puntajes y las variables de criterio escogidas. Se estima que los estudios de regresión son apropiados y que los indicadores generales de asociación deben ser suplementados con los índices de asociación en cada uno de los rangos de puntajes de la prueba²⁹. La prueba de selección utilizada hasta el año 2002 (PAA) logra indicadores de predictibilidad similares a los obtenidos en las baterías de selección en EE.UU. Las nuevas evaluaciones debieran dar garantías de que al menos logran índices similares de

²⁸ Véase Estándar 1.7 en AERA, APA, NCME, *Standards for Educational and Psychological Testing* (1999).

²⁹ Véase Estándar 1.15, en *ibídem*.

<i>Estudios de sesgo</i> <i>Estudio de las consecuencias de las pruebas</i>	predictibilidad antes de entrar a reemplazarla. Al igual que en otros países, se esperaría que se condujeran estudios piloto antes de las aplicaciones definitivas para cerciorarse de que las inferencias que se harán tienen respaldo empírico. Por otra parte, los estándares exigen que se expliciten las características técnicas de los estudios de validación predictiva y que queden claramente documentados los tipos de ajustes estadísticos, por restricción de rango, realizados en las estimaciones³⁰. Nada de ello aparece en el proyecto FONDEF. El proyecto FONDEF especifica que en la etapa experimental y después de la primera aplicación de la prueba se realizarán los estudios de sesgo correspondientes de acuerdo a la técnica DIF. El Informe de la Comisión, por su parte, asevera que las consecuencias asociadas a las pruebas de ingreso a la universidad actúan como un incentivo para el estudio y que por tanto el contenido de la evaluaciones debiera alinearse con el currículum para mejorar el rendimiento de los alumnos en la educación media. Según los estándares de la AERA, cuando explícitamente se sugiere que el uso de un test producirá determinados resultados, se debiera entregar la evidencia teórica y empírica que permita sustentar dicha afirmación³¹. En el proyecto no se señala cómo se validará empíricamente la hipótesis planteada y no se entregan argumentos sólidos acerca de cómo la nueva prueba mejorará el rendimiento de alumnos competentes que ya contaban con incentivos fuertes para el estudio, dado que las notas de educación media se consideran al momento de ingresar a la universidad.
Consideraciones de seguridad	El proyecto estima que un tercio de las preguntas de cada prueba se repetirá en el siguiente año para efectos de comparación de las pruebas. Los autores afirman que “se intentará retener alrededor de un tercio de las preguntas para hacer la

³⁰ Véase Estándar 1.18, en ibídem.

³¹ Véase Estándar 1.23, en ibídem.

	comparabilidad interanual”³². Por otra parte, se ha considerado incluir también en las pruebas definitivas las preguntas experimentales de las pruebas de los años siguientes. Dado que en el país existe una industria activa de preuniversitarios y un interés grande de los establecimientos educacionales por conocer el tipo de preguntas que se harán en los años siguientes, cabe preguntarse cómo controlarán la filtración de preguntas desde los alumnos que dan las pruebas hacia estas instituciones. El proyecto FONDEF no hace referencias al respecto.
Documentación del proceso	En el proyecto no se consigna presupuesto ni tiempo para documentar el proceso de construcción de la prueba. De hecho los únicos documentos públicos conocidos son el Informe de la Comisión y el proyecto FONDEF. Una prueba que tiene consecuencias altas para más de la mitad de los alumnos de cada generación debiera hacer pública y transparente la fundamentación y procedimientos de cada uno de sus pasos³³.
Validación de la prueba ante el público	El proyecto FONDEF de Bravo y Manzi consideró la publicación de preguntas de apoyo pedagógico que difundieran la naturaleza de las pruebas. Sin embargo, en el cronograma se observa un desfase importante entre la confección de los modelos de pruebas que se darían a conocer al público y el análisis experimental de las preguntas que determinarían la extensión total de la prueba y el nivel de dificultad de la misma. Por tanto, los modelos de pruebas que se darían a conocer al público no necesariamente coincidirían con la prueba real. Esto representa un problema para la validación de la prueba ante el público porque no permite juzgar públicamente su pertinencia. Por otra parte, la falta de documentación del proceso de construcción de la prueba no ayuda a que los actores relevantes avalen la prueba. Por el contrario, crea suspicacias que le restan validez.

³² Véase Bravo y Manzi, “Reformulación de las Pruebas de Selección a la Educación Superior” (proyecto FONDEF) (2001), p. 42.

³³ Véase Estándar 6.1 al 6.7, en AERA, APA, NCME, *Standards for Educational and Psychological Testing* (1999).

2.3. Pruebas de Selección Universitaria (PSU) 2003

¿Qué posibilidad existe de que, en el período que se ha fijado para la construcción de las nuevas pruebas de admisión, los equipos responsables puedan elaborar una prueba que cumpla con los estándares mínimos que garanticen un proceso de selección justo, válido y confiable?

Para analizar el problema es importante considerar el cronograma y el detalle de los principales hitos de la construcción de la nueva prueba. El 29 de agosto del año 2002 se anuncia que la generación que egrese de educación media el 2003 se seleccionaría con una prueba de admisión transitoria. El Consejo de Rectores decide que su construcción será coordinada por un Comité Técnico Asesor nombrado por el Consejo Directivo para las Pruebas de Selección y Actividades de Admisión a la Universidad, entidad que se constituyó el 4 de septiembre del año 2002, para monitorear el desarrollo de las pruebas. Dicho Consejo Directivo emite un comunicado el 15 de noviembre del año 2002, que define sucintamente las características de las nuevas pruebas³⁴. Se acuerda que los postulantes tendrán que rendir tres pruebas. De estas tres, dos serán obligatorias: lenguaje y matemática. Para la tercera habrá que optar entre una prueba de historia y geografía y otra de ciencias. Esta última estará dividida, a su vez, en una primera sección común referida a los contenidos curriculares de biología, física y química de I y II año medio, y una segunda sección de tres módulos, en la cual los alumnos tendrán que escoger uno de ellos. Estos módulos corresponderían a los contenidos de biología, física o química de III y IV medio.

En el mismo documento se fijan además las ponderaciones mínimas para cada uno de los factores considerados en el proceso de selección. El promedio de notas de enseñanza media no puede contar menos de un 20% y las pruebas de ciencias o de historia y geografía; de lenguaje y matemática no puede contabilizarse menos de un 10%. Ponderar en mayor proporción que los mínimos señalados queda en manos exclusivas de cada universidad, de acuerdo a los propósitos de selección que ellas estimen conveniente.

El comunicado del 15 de noviembre (2002) también afirma que los contenidos curriculares de la educación media a ser considerados como referentes de las pruebas quedaron fijados por el Comité Técnico Asesor del Consejo Directivo, acogiendo la proposición concordada por la Mesa

³⁴ Consejo de Rectores; Consejo Directivo para las Nuevas Pruebas de Selección y Actividades de Admisión a la Universidad. "Sobre las Pruebas de Selección a la Enseñanza Universitaria 2003", viernes 15 de noviembre de 2002.

de Trabajo de los Sostenedores, Colegio de Profesores y Ministerio de Educación.

El 17 de noviembre se publica el listado de los contenidos curriculares de enseñanza media que se tomarán como base para la construcción de la prueba transitoria. Corresponde a una transcripción precisa de los contenidos mínimos del marco curricular nacional de educación media, exceptuando aquellos contenidos que la Mesa de Trabajo consideró que no habían sido cubiertos por una proporción amplia de los alumnos que rendirían la prueba³⁵.

El 28 de enero de 2003 el Consejo Técnico Asesor menciona cuáles serán los ejes temáticos de cada una de las pruebas y las habilidades intelectuales que se exigirán, exceptuando la de ciencias, que queda sin mayores precisiones. Se entrega también el número total de preguntas y el tiempo de cada evaluación. Finalmente, se da a conocer el calendario de actividades hasta la fecha de aplicación de la prueba. En la calendarización se incluyen los pasos que involucran a los alumnos y universidades, pero no se incluyen las etapas técnicas del desarrollo de las pruebas.

En marzo de 2003 se publica un folleto con los contenidos, tabla de especificaciones, pero sin los pesos asignados a cada sección, y un número reducido de preguntas de las nuevas pruebas. En dichos folletos se entrega una descripción breve de la naturaleza de la prueba. En esa misma fecha, las universidades del Consejo de Rectores publican la nómina de carreras con las pruebas optativas y las ponderaciones finales que se pedirán en cada caso.

Entre abril y mayo de 2003 se difundirán los folletos de cada una de las pruebas con ejemplos de cada sección. Los documentos oficiales dejan entrever que se tratará de un muestreo de preguntas más que facsímiles de pruebas completas. La fecha de rendición de las pruebas queda fijada para mediados de diciembre. Este cronograma supone que las pruebas quedarán prácticamente definidas entre el 4 de septiembre de 2002, fecha en que se conforma el Comité Asesor, y mayo de 2003, en que se dan a conocer los facsímiles.

En declaraciones no oficiales, se han ido estipulando aspectos que no aparecen definidos en los comunicados anteriores. La presidenta del Comité Técnico Asesor afirma que aun cuando se deben rendir obligatoriamente tres de las cuatro pruebas, los alumnos podrían darlas todas. En caso de que las carreras exigieran indistintamente las pruebas de ciencias o la de historia y geografía, el estudiante podrá postular con la que obtuvo mejor puntaje. En relación de la prueba de ciencias, las autoridades universitarias

³⁵ Referente Curricular de Pruebas de Selección Universitaria. www.mineduc.cl

establecieron que la prueba tendrá un puntaje único pese a contar con dos partes, una de las cuales no es la misma para todos los individuos. Es decir el puntaje final de la prueba de ciencias no distinguirá si el alumno ha rendido la segunda parte en biología, física o química. Esto implica que un estudiante, por ejemplo, podrá postular a medicina y a ingeniería dando la prueba común de ciencias con la optativa de biología. Por otra parte, aclaran que el margen de libertad que se le entrega a cada universidad para ponderar las distintas pruebas debe emplearse dentro de las cuatro pruebas establecidas, lo cual excluye la posibilidad de considerar pruebas específicas anexas en la ponderación total.

Hasta aquí la información oficial disponible. Un análisis de la viabilidad de la nueva prueba necesariamente pasa por el campo de las suposiciones, ya que los datos que se han dado a conocer son de carácter amplio y no incluyen las descripciones técnicas de la construcción de los nuevos instrumentos.

A continuación se explorará el escenario de las Pruebas de Selección Universitaria (PSU) 2003, basándose en el supuesto de que éstas combinan elementos de la Prueba de Aptitud Académica (PAA), de las Pruebas de Conocimientos Específicos y de los módulos experimentales que posiblemente ha elaborado y testeado previamente el DEMRE³⁶ en el contexto de investigaciones internas para el mejoramiento de las pruebas de admisión y secciones de las pruebas realizadas en el marco del proyecto SIES.

Se descartan otros escenarios dada la limitación de tiempo que impuso el Consejo de Rectores. La posibilidad de incluir secciones de pruebas de admisión de probada calidad que no han sido testeadas en Chile quedó seriamente limitada, al menos para la evaluación 2003. Por ejemplo, no se podría contemplar la adaptación de secciones del ACT de ciencias³⁷, aun cuando pedagógicamente es una prueba que recoge lo crucial que deben aprender los alumnos en educación media y constituye un buen indicador de éxito en la universidad. En efecto, el ACT de ciencias podría ser un modelo a seguir en esa área, ya que el tipo de preguntas que emplea permite evaluar razonamiento científico y aplicación de conceptos centrales, sin recurrir a la memorización innecesaria de definiciones mecánicas³⁸. Sin

³⁶ DEMRE (Departamento de Evaluación, Medición y Registro Educacional), institución encargada del proceso de admisión a la universidad, dependiente de la Universidad de Chile.

³⁷ El ACT (American College Testing Program) es una de las pruebas de selección universitarias de Estados Unidos. En ese país, es la segunda en importancia según el número de alumnos que la rinde.

³⁸ Este punto quedó destacado en el Informe de la Comisión de Ciencias del CEP, presidida por Sergio Hojman, en octubre de 2002. Véase Comisión de Ciencias, "Prueba de Ciencias, Críticas y Propuestas" (2002).

embargo, dado el cronograma que se ha fijado, ya es tarde para intentar adaptar secciones del ACT, porque en menos de un año no se podría entrenar a los redactores de ítems en el nuevo formato, adaptar los ítems a los contenidos del currículum chileno, evaluar empíricamente las preguntas para estudiar su capacidad discriminativa y su comportamiento psicométrico, analizar con jueces expertos la validez de contenido y finalmente conducir los estudios de validez predictiva. En este plazo, también resulta imposible familiarizar a los profesores y alumnos con el estilo de las preguntas.

El mismo fenómeno ocurre con la posibilidad de incluir secciones nuevas desarrolladas en el país, aunque en este caso la situación se agrava aún más. Las innovaciones requieren de la elaboración y validación de un marco que avale teórica y empíricamente la decisión de incluir un tipo de pregunta. Cuando se adaptan pruebas, sólo debe estudiarse la validez de generalizar el constructo a otros contextos, lo cual requiere menos tiempo.

En suma, a continuación se analizará la PSU (Pruebas de Selección Universitaria) 2003, suponiendo que ella quedará conformada por una mezcla de preguntas extraídas de la Prueba de Aptitud Académica, de las Pruebas de Conocimientos Específicos, del SIES y por nuevas preguntas elaboradas en el contexto de las investigaciones del DEMRE.

(Pruebas de Selección Universitaria (PSU) 2003)

Etapas de especificación de la prueba	Las definiciones que debe realizar el Comité Técnico Asesor están fuertemente condicionadas por las restricciones técnicas que impone un calendario tan ajustado como el que se le ha impuesto a la prueba. Más que elegir las características idóneas de cada prueba, el Comité tiene que escoger de un menú de secciones y preguntas que han sido testeadas anteriormente. El Comité también parte con un pie forzado, pues el número de pruebas ha quedado limitado <i>a priori</i>, a pesar de la evidencia empírica que destaca el valor de tener una combinación de pruebas generales y específicas. En efecto, el mandato del Consejo de Rectores impide agregar pruebas a las tres obligatorias, con lo cual se elimina la posibilidad de incorporar pruebas específicas que han demostrado un alto valor predictivo³⁹.
--	--

³⁹ Véanse Vial y Soto, “¿Predice la PAA el Rendimiento o el Éxito en la Universidad? (2002), y Fischer y Repetto, “Método de Selección y Resultados Académicos” (2002).

<i>Propósito de la prueba</i>	Por otra parte, la prescripción de utilizar el currículum de educación media como referente de la prueba también podría representar una cortapisa para la construcción de la prueba, como se explicará a continuación. El Comité Técnico a cargo de las nuevas pruebas asegura que su objetivo principal es seleccionar a los mejores estudiantes para la educación superior y que no es una meta evaluar la educación media⁴⁰. En el documento de difusión de los contenidos, tabla de especificaciones y muestreo de preguntas aparecido en marzo de 2003, se enfatiza que la selección es uno de los objetivos principales de la prueba. Desde luego, allí se bautiza la nueva prueba con el título de “Prueba de Selección a la Universidad” y se asevera que la nueva prueba “responde a los requerimientos de una mayor alineación con los Programas de Estudio vigentes en la Enseñanza Media de cada uno de los tests que la componen, sin perder su característica principal de ser pruebas de selección”⁴¹. Sin embargo, este sucinto documento no alcanza a constituirse en el marco de fundamentación de la prueba, ya que no entra en detalles, delimitaciones de conceptos ni provee la evidencia teórica y científica que avale el contenido y la orientación del instrumento. Por lo tanto, no debiera tomarse como la última palabra. En cuanto a los propósitos de la prueba y los constructos teóricos que guiarán su elaboración, el documento del 15 de noviembre del Consejo de Rectores no se pronuncia al respecto. Solamente estipula que los contenidos de la educación media deben ser utilizados como referentes de la prueba. Esta afirmación da pie a distintas interpretaciones. Por una parte, utilizar los contenidos de educación media como referente podría significar que los evaluadores pueden escoger, entre todos los contenidos del ciclo, aquellos cuyo dominio es el indicador de un buen desempeño en la universidad.
--------------------------------------	---

⁴⁰ *El Mercurio*, miércoles 29 de enero 2003.

⁴¹ Consejo de Rectores; Departamento de Evaluación, Medición y Registro Educacional, “Pruebas de Selección Universitaria: Proceso de Admisión 2004”, *La Nación*, marzo 2003.

O bien que, al construir pruebas centradas en la evaluación de destrezas y habilidades requeridas para cursar estudios superiores, los evaluadores tendrán que utilizar los contenidos del currículum de educación media que mejor se presten para ello. Estas dos interpretaciones avalarían la posición de que las pruebas de admisión 2003 tienen como propósito fundamental seleccionar a los estudiantes que tienen mayores posibilidades de éxito en la universidad.

En cambio, se puede desdibujar el objetivo de seleccionar a los que tienen mayores probabilidades de éxito en la universidad si utilizar el currículum como referente significa un vuelco a priori hacia pruebas centradas en el dominio del currículum, es decir a la evaluación aleatoria de sus contenidos. El coordinador nacional de la Unidad de Currículo y Evaluación, del Ministerio de Educación, miembro del Consejo Técnico Asesor, asevera que “la esencia del cambio de las pruebas de selección, decidido por el Consejo de Rectores, es el giro de las mismas hacia una mayor cobertura curricular. La más importante ventaja de las pruebas de selección basadas en el currículo consiste en un incentivo claro y fuerte para el trabajo de alumnos y profesores en la enseñanza media, valorizando el quehacer de ésta y su profesorado, y produciendo egresados mejor preparados”⁴². La pregunta es si todos los contenidos de la educación media actual convergen con los requerimientos de la educación universitaria y si esto debiera ser así, dado que este ciclo puede ser un fin en sí mismo y no una preparación para la universidad.

Estos planos deben aclararse, ya que al evaluar la validez de contenido de la prueba, los jueces tendrán que decidir si la estructura de la batería de selección y las preguntas son fieles al objetivo de seleccionar a los alumnos a la universidad o si, además de seleccionar, logran ser un fiel referente de los programas educativos, con lo cual se acercarían al objetivo de ser un incentivo claro y fuerte para el dominio del currículum de la educación media.

⁴² C. Cox, “Nuevas Pruebas de Selección Universitaria: En Aguas Más Calmas” (2002).

<i>Fundamentación</i>	Si se emplea una combinación de secciones que provienen de distintas pruebas, el marco conceptual debiera fundamentar por qué esa amalgama particular responderá al propósito de la prueba. El marco de la prueba debe definir claramente qué hay detrás del conjunto de preguntas que se emplearán. Luego, debe aportar evidencia teórica que indique por qué la batería de pruebas escogidas y el tipo de dominios y habilidades a evaluar predecirán el éxito en la universidad. Junto con ello, en la etapa inicial del desarrollo de una prueba se debe aportar evidencia empírica que demuestre por la vía de la experiencia comparada que esto será así. En el caso de estas pruebas, en que el marco teórico se tendrá que construir post hoc a las decisiones que tomó el Consejo de Rectores, será especialmente difícil demostrar, por ejemplo, que buenos resultados en la prueba de ciencias con el módulo de biología resultará predictiva del éxito en la carrera de ingeniería civil. Asimismo, debe fundamentarse por qué el dominio de ciertos contenidos puntuales del currículum de lenguaje incluidos en la prueba constituirán un indicador de éxito en la mayoría de las carreras⁴³. Lo mismo ocurre con los contenidos de la prueba de matemática. Aquí debe quedar justificado cuáles contenidos y destrezas son relevantes de evaluar y por qué se evaluarán de una manera u otra. Debe quedar claramente descrito cuál es la manera de preguntar que respeta el constructo de la prueba. A modo de ejemplo, para predecir éxito en las carreras matemáticas podría ser importante la evaluación de los teoremas y posiblemente es más predictivo exigir su aplicación que su memorización. En este acápite se debiera dar cuenta también de las razones por las cuales se eliminaron las Pruebas Específicas y justificar con evidencia que las pruebas que las reemplazarán aportarán más a la predictibilidad.
------------------------------	--

⁴³ En la prueba de lenguaje se incluyen preguntas relacionadas con literatura. Esto exige una fundamentación distinta a la que comúnmente se da cuando se aplican pruebas verbales generales, ya que normalmente las secciones de literatura se incluyen en las pruebas específicas de lenguaje, las cuales sólo se utilizan para seleccionar alumnos para carreras afines.

<i>Definición de los dominios a evaluar</i>	Por el momento no se cuenta con ningún documento que reúna las condiciones anteriores. A partir de las “Tablas de Especificaciones de las Pruebas de Selección Universitaria 2003” y del “Temario de las Pruebas de Selección Universitaria 2003”⁴⁴, dos equipos paralelos, trabajando independientemente, no podrían construir pruebas similares, tal como lo sugieren los estándares. En primer lugar, aún no se ha definido el peso que se le asignará a cada habilidad y eje temático. Pero si dejamos este aspecto de lado, tampoco se podría lograr este objetivo. El temario se refiere a los contenidos del marco curricular de educación media, y éste no precisa bien la profundidad con que se debe tratar cada uno de los tópicos allí presentados⁴⁵. El marco enuncia los temas pero no acota el nivel de detalles que comprende. En el caso de la prueba de lenguaje, las incógnitas son mayores, pues nunca antes se había evaluado el conocimiento de “unidades conceptuales relevantes” de la asignatura y tampoco aspectos de la teoría del discurso. Por lo tanto, los encargados de la prueba debieran definir mejor los objetivos de logro para cada contenido y delimitar claramente el alcance de cada uno de ellos. Sólo de este modo los alumnos sabrán a qué atenerse al momento de estudiar la prueba. No basta con el currículo nacional, porque éste expone los contenidos a grandes rasgos, lo cual obedece al mandato de la Ley Orgánica Constitucional de Educación de dar la libertad a cada establecimiento para elaborar sus propios planes y programas. Por esta misma razón, la concreción del currículo en los planes y programas de estudio del Ministerio no puede servir de guía porque no tiene carácter obligatorio para todos los establecimientos. Otro tanto sucede con los textos de estudio, ya que éstos no son únicos ni obligatorios.
--	--

⁴⁴ Consejo de Rectores, Documento N°2 y N°3, enero 2003.

⁴⁵ Se podría argumentar que se daba el mismo caso en las Pruebas de Conocimientos Específicos. Sin embargo, la historia acumulada de las pruebas, recogida en los facsímiles oficiales, lograba acotar claramente, vía ejemplo, la profundidad y alcance de cada uno de los contenidos a evaluar.

<i>Formato de la prueba</i>	El número de preguntas y tiempo de duración de la prueba quedó establecido el 23 de enero, posiblemente antes de definir el tipo de ítems que se emplearía. Es probable que esto se haya hecho para despejar la incertidumbre de los estudiantes ante la vecindad de la prueba. Normalmente esto no sucede así, ya que el número de preguntas responde a una conjunción de factores que se ponderan y se van afinando durante el proceso de elaboración de los ítems. Desde luego, la definición del constructo de la prueba debe anteceder a la determinación del tiempo total de la prueba y al número de preguntas. Por ejemplo, si las preguntas requieren razonamiento avanzado, normalmente exigirán más tiempo para contestarlas, lo cual resulta determinante para el número final de preguntas. En el documento de marzo de 2003 se da a conocer el formato de las preguntas y se espera que el peso de cada sección y el modelo definitivo de las pruebas quede definido entre abril y mayo.
<i>Características psicométricas</i>	Una de las primeras decisiones que deben adoptarse es si el cálculo de los puntajes de los alumnos se hará a través de la Teoría Clásica (TC) o de la Teoría de Respuesta al Ítem (IRT). Si bien ambos análisis son complementarios, tienen fuertes diferencias a la hora de determinar el puntaje: en Teoría Clásica, cada pregunta vale lo mismo, en IRT se ponderan por su nivel de dificultad. Para los alumnos es importante conocer y entender la forma en que se obtendrá su puntaje. Un estudiante que se prepara para la prueba puntuada por IRT debe considerar que el cálculo de los puntos no es lineal. Por lo tanto, su puntaje total no corresponderá a la suma de repuestas correctas (menos las incorrectas) sino que es un puntaje en el que un mismo número de respuestas correctas puede representar un puntaje distinto, dependiendo del nivel de dificultad de las mismas. También tendrá que informarse que no se descuentan preguntas como se hace con la Teoría Clásica. La decisión de utilizar IRT debe comunicarse junto con la publicación de los facsímiles. En el caso de la prueba

de admisión 2003, ésta es una decisión que estará condicionada por la factibilidad de aplicar IRT⁴⁶. Una prueba de selección a la universidad requiere características psicométricas especiales. Debe permitir una jerarquización fina de los estudiantes en los segmentos de corte de cada carrera, para evitar márgenes de error significativos. Como en Chile los puntos de corte son distintos en cada una de las carreras y universidades, se esperaría que las pruebas discriminaran bien en un amplio rango de habilidades. Al menos de la media hacia arriba (niveles de habilidades sobre el promedio) que es el tramo del que se elegirán a los estudiantes. Todo esto implica que la especificación de la distribución de los ítems, según nivel de dificultad (si se emplea Teoría Clásica), o la curva de información (si se emplea Teoría de Respuesta al Ítem), debiera concentrarse en el rango superior de competencias. Si bien es posible especificarlo en el plano teórico, será difícil utilizar esta especificación para construir la prueba. La principal razón se debe a que se ensamblará una prueba a partir de secciones, testeadas experimentalmente en contextos diferentes, sin un diseño que permita una comparación de los resultados. Así, los niveles de dificultad (tanto en Teoría Clásica como IRT) no serán comparables. Los constructores de la nueva prueba requerirán realizar supuestos arriesgados para poder asumir equivalencia. Este mismo problema se enfrenta con los demás descriptores de los ítems: nivel de discriminación, y en el caso del modelo IRT usado en el proyecto FONDEF se debe agregar el que corresponde “a la adivinación”; niveles de homogeneidad de los ítems si se trata de la teoría clásica o unidimensionalidad en el caso de IRT. En la batería de pruebas anterior se contaba con tests generales y específicos. Estos últimos facilitaban la tarea de discriminación fina en el segmento superior de desempeño. Con la eliminación de ellos en la nue-

⁴⁶ La posibilidad de usar IRT en las nuevas pruebas sólo se podrá definir tras la aplicación definitiva de la PSU 2003. No sería factible decidirlo antes porque la nueva prueba posiblemente se habrá construido a partir de dos pruebas que siguen modelos distintos. Si la prueba PSU rendida en diciembre de 2003 se comportara de acuerdo a las especificaciones previamente definidas, entonces se podrían calcular los puntajes con IRT. Sin embargo, esto sería éticamente reprochable porque los alumnos no conocerán las reglas del juego sino hasta después de rendida la prueba.

	va prueba, se cuenta con menos preguntas. Las especificaciones debieran definir cómo compensarán esta pérdida de información. La comparabilidad interanual de las pruebas no debiera buscarse en esta etapa porque condicionaría las futuras evaluaciones. Estas tendrían que seguir un formato similar a la prueba 2003, lo cual no sería recomendable dadas las particulares circunstancias de esta prueba. Además se ha reconocido que es una prueba de transición. Por lo tanto, a nuestro juicio, no debieran especificarse los requerimientos de <i>equating</i> para el instrumento del 2003.
Construcción de los ítems	Los ítems fueron contruidos con anterioridad a la especificación de las pruebas. Sin embargo, para validar su utilización en una prueba con especificaciones distintas a las cuales les dieron origen, debieran someterse al juicio de los expertos para verificar su adecuación a los nuevos requerimientos. Procedimiento que debiera quedar debidamente documentado. Este trabajo correspondería a la etapa de revisión inicial de los ítems y a la de evaluación preliminar. El estudio piloto de las preguntas ensambladas en versiones similares a las finales probablemente no podrá realizarse a menos que se haga una aplicación de campo durante el año 2003. Hecho que no se ha anunciado. Este estudio es el que permite determinar con un alto grado de precisión el comportamiento psicométrico de las preguntas. Por lo tanto, si no se realiza, la conformación de las pruebas finales se tendrá que hacer con datos aproximados obtenidos de los estudios piloto que se hicieron en forma independiente para cada prueba original de donde se sacaron los ítems para esta nueva versión. Como se dijo, los criterios técnicos indican que los parámetros de las preguntas pueden variar cuando las preguntas se cambian de contexto, lo cual indicaría que los datos obtenidos en las condiciones originales no son directamente extrapolables en versiones distintas. Esto implica que los indicadores de dificultad y parámetros psicométricos de las preguntas pueden variar al momento de la aplicación definitiva de la prueba. En el peor de los casos las pruebas po-

	drían no discriminar bien en algunos segmentos. Por ejemplo, como se estimaron los niveles de dificultad antes de que los alumnos se prepararan para las pruebas, podría suceder que una vez que éstos estudien, las preguntas que consideraban difíciles ya no lo sean más. Con ello existiría el riesgo de que una proporción alta de alumnos se agolpara en los puntajes superiores.
Estudios de confiabilidad y validez <i>Estudios de confiabilidad</i> <i>Validez de constructo y de contenido</i> <i>Validez predictiva</i> <i>Estudios de sesgo</i>	Si la nueva prueba mezcla ítems procedentes de tests de naturaleza distinta, probados en condiciones diferentes, es difícil que se puedan realizar estudios de confiabilidad interna de la prueba completa antes de su aplicación definitiva. En este sentido no se cumpliría con los estándares de la AERA que exigen que las pruebas cuenten con un margen de confiabilidad al aplicarse. Antes de la aplicación de la prueba debiera darse a conocer la nómina de expertos que revisarán las pruebas y los informes que emitan. Previo a esto, tendría que formularse claramente la fundamentación de la prueba para que los jueces pudieran tener los criterios definidos para emitir sus juicios. La tarea de este grupo será muy difícil porque en el caso de que su evaluación sea negativa, hay poco tiempo para hacer correcciones. ¿Qué comisión se atrevería a invalidar una prueba? ¿Hay una prueba alternativa en caso de que esta no fuera viable? Dados los tiempos fijados, la nueva prueba se aplicará sin evidencia empírica de su validez predictiva. Para tenerla se requieren al menos dos años desde una aplicación piloto. Como afirma el estudio de Fisher y Repetto (2002), se reemplazará una prueba que tiene validez predictiva por una que no ha demostrado tenerla. Los estudios de apariencia de sesgo de los ítems pueden realizarse durante el análisis preliminar de los ítems. Sin embargo, probablemente la aplicación del análisis de DIF se podrá realizar sola-

<i>Estudio de las consecuencias de las pruebas</i>	mente a las pruebas originales de donde se sacarán los ítems, ya que los estudios experimentales se realizaron por separado para cada una de ellas, con muestras independientes. El análisis de sesgo de la prueba total sólo se podrá conducir una vez que esta se aplique definitivamente, lo que no es el ideal. Si la nueva prueba afirma que los alumnos de educación media aumentarán su dedicación al estudio y con ello mejorará el rendimiento académico en este ciclo, habrá que recoger evidencia empírica al respecto una vez aplicados los tests. Ésta es una tarea que tendría que demostrar que hay un aumento del rendimiento en un período determinado y que éste se puede atribuir a la aplicación de las pruebas y no a otros factores. Por otra parte, también hay que recoger evidencia de que la prueba no distorsionará el estudio del currículum en este ciclo. Por ejemplo, si se pregunta en la prueba por conceptos literarios, habrá que investigar si este hecho conduce a que se deje de leer el mínimo de seis obras literarias por año para dar espacio a la ejercitación descontextualizada de dichos conceptos.
Consideraciones de seguridad	Si la nueva prueba no se diseña para hacer sus puntajes comparables de un año a otro y no se experimentan los ítems de la prueba del 2004 en la evaluación del 2003, las condiciones de seguridad no tendrían por qué variar respecto de años anteriores. En este sentido el DEMRE, el organismo que administrará las pruebas, ha dado garantías de seguridad en las últimas evaluaciones.
Validación de la prueba ante el público	Dada la precariedad de las nuevas pruebas, su validación ante el público será difícil. La claridad y transparencia serán cruciales para ganar el favor del público. En este sentido, la publicación de modelos completos de las pruebas y un análisis favorable de las mismas por jueces externos al circuito del Consejo de Rectores ayudarían a atenuar la desconfianza inicial. Sin embargo, un test con efectos retroactivos como éste, difícilmente será percibido como justo por alumnos y profesores. La reforma curricular

	sobre la cual se basa la prueba se ha prestado para distintas interpretaciones lícitas; esta nueva prueba favorecerá necesariamente una línea y no otras. Lo más equitativo habría sido aplicar la prueba una vez que los estudiantes de I año medio, en antecedente de la misma, terminaran el IV medio.
--	---

Claramente, el procedimiento seguido en la construcción de esta nueva prueba de admisión no ha sido el regular. Del análisis precedente se puede concluir que las posibilidades de cumplir con los estándares requeridos para la construcción de pruebas de altas consecuencias antes de su aplicación son mínimas.

Es probable que en países con una tradición más vasta en la construcción de pruebas, con una masa crítica de evaluadores formados que pueden ejercer el rol de contraparte en el proceso de la elaboración de los tests, no se habría aceptado la puesta en marcha de la Prueba de Selección Universitaria 2003 para el proceso de admisión 2004.

Sin embargo, dado que la decisión de aplicarla está tomada y que echar pie atrás induciría a una confusión mayor, los organismos competentes debieran intentar minimizar las carencias. En primer lugar, avanzar en la definición de la fundamentación de la prueba para permitir que jueces expertos validen los contenidos de la prueba. Luego se puede proceder a demostrar que ciertas secciones de la prueba tienen un marco de validación mínimo, por la vía de la evidencia comparada con otras pruebas que las contienen y que han sido validadas empíricamente. En tercer lugar, comprometer y calendarizar los futuros estudios de validación que se realizarán una vez que se apliquen las pruebas y detallar los procedimientos que se seguirán si éstos no son favorables. En cuarto lugar, consignar los procedimientos seguidos en la confección de las pruebas para facilitar los estudios de validez y asegurar transparencia total. Por último, es fundamental que los responsables garanticen que en el futuro nuestras pruebas se alinearán con los estándares internacionales.

La Prueba de Admisión 2003 no debiera constituirse en un marco forzado para las evaluaciones de los años subsiguientes. Sólo debiera ser un antecedente más en el estudio de la batería idónea para seleccionar a los mejores alumnos para la universidad. En este sentido, habría que reabrir el debate y los estudios acerca de la conveniencia de eliminar las pruebas específicas, investigar la utilidad de adaptar nuevas pruebas o secciones de pruebas extranjeras, buscar evidencia empírica que avalara o revocara la decisión de exigir pruebas de asignaturas no afines a las carreras o líneas de

estudio. Y por qué no, también se debiera explorar la posibilidad de contratar los estudios de validación a instituciones de prestigio como el Educational Testing Service (ETS), o investigar si sería mejor volver a la batería de pruebas anteriores incorporando las modificaciones que se han hecho a las pruebas equivalentes en EE.UU. Abrirse a estas posibilidades permitirá contar con pruebas justas, confiables y pertinentes.

BIBLIOGRAFÍA

- American Educational Research Association (AERA), American Psychological Association (APA), National Council on Measurement in Education (NCME). *Standards for Educational and Psychological Testing*. AERA, 1999.
- Barret, P. "How to Review a Psychological Test Instrument. University of Liverpool", 1999. www.liv.ac.uk/pbarret/paulhome.htm.
- Beyer, H. "Las Nuevas Pruebas de Ingreso a la Universidad". *Administración y Economía*, 48, Revista de la Facultad de Ciencias Económicas y Administrativas, PUC, 2002.
- Bravo, I. y J. Manzi. "Reformulación de las Pruebas de Selección a la Educación Superior". Proyecto FONDEF (Octavo Concurso Nacional de Proyectos de Investigación y Desarrollo), 2000.
- Camara, W. J. y G. Echternacht. "The SAT I and High School Grades: Utility in Predicting Success in College". Research Notes RN-10, The College Board, julio 2000.
- Comisión de Ciencias del CEP. "Prueba de Ciencias, Críticas y Propuestas". *Estudios Públicos* 88 (2002).
- Comisión Nuevo Currículum de la Enseñanza Media y Pruebas del Sistema de Admisión a la Educación Superior. "Informe de la Comisión Nuevo Currículum de la Enseñanza Media y Pruebas del Sistema de Admisión a la Educación Superior". Santiago, noviembre 2000.
- Consejo de Rectores; Consejo Directivo para las Nuevas Pruebas de Selección y Actividades de Admisión a la Universidad. "Sobre las Pruebas de Selección a la Enseñanza Universitaria 2003". Viernes 15 de noviembre de 2002.
- Consejo de Rectores. Documentos N° 1, N° 2 y N° 3. Enero 2003. www.consejodirectores.cl
- Consejo de Rectores; Departamento de Evaluación, Medición y Registro Educacional. "Pruebas de Selección Universitaria: Proceso de Admisión 2004". Diario *La Nación*, marzo, 2003.
- Cox, C. "Nuevas Pruebas de Selección Universitaria: En Aguas Más Calmas". *Revista de Educación*, 301. Ministerio de Educación, diciembre 2002.
- Departamento de Evaluación, Medición y Registro Educacional (DEMRE). "Calendario Básico de Trabajo y Publicaciones". Rendición de Pruebas de Selección Universitaria 2003. Documento N° 1. Universidad de Chile. Enero 2003.
- Departamento de Evaluación, Medición y Registro Educacional (DEMRE). "Tabla de Especificaciones de las Pruebas Universitaria 2003". Documento N° 2. Universidad de Chile. Enero 2003.
- Departamento de Evaluación, Medición y Registro Educacional (DEMRE). "Temario de las Pruebas de Selección Universitaria 2003". Documento N° 3. Universidad de Chile. Enero 2003.

- Donoso, G.; M. A. Bocchieri, E. Ávila y otros. *El Sistema de Admisión: Orígenes y Evolución. Resultados del Proceso de Admisión 1999*. Universidad de Chile, DEMRE. 1999.
- Educational Testing Service (ETS). *Standards for Quality and Fairness*. ETS, 2000.
- Fischer, R. y A. Repetto. “Método de Selección y Resultados Académicos”, Escuela de Ingeniería, Universidad de Chile”. Universidad de Chile, noviembre 2002.
- Hambleton, R. K.; H. Swaminathan y H. J. Rogers. *Fundamentals of Item Response Theory*. Newbury Park: Sage. 1991.
- Le Foulon, C. “Reformulación del Sistema de Selección a la Educación Superior. Antecedentes para la Discusión”. Centro de Estudios Públicos, *Serie Documentos de Trabajo*, 334, 2002.
- Le Foulon, C. y F. Dussaillant. “Desarrollo de Pruebas Estandarizadas”. Mimeo, Centro de Estudios Públicos. 2002.
- Ramist, L.; C. Lewis y L. McCamley-Jenkins. “Using Achievement Test/Sat II: Subject Tests to Demonstrate Achievement and Predict College Grades: Sex, Language, Ethnic, and Parental Education Groups”. College Board Research Report N° 2001-5, 2001.
- Referente Curricular de Pruebas de Selección Universitaria. www.mineduc.cl
- Shepard, L. “Evaluating test validity”. *Review of Research in Education*, 19. 1993.
- Rosas, R.; M. P. Flotts, y C. Saragoni. “Modelo de Representación del Conocimiento para las Nuevas Pruebas de Selección para el Ingreso a las Universidades Chilenas”, *Psykhé*, vol. 11 N° 1, Pontificia Universidad Católica de Chile, 2002.
- Tinkelman, S. “Planning the Objective Test”. En Thorndike, R. L. (ed). *Educational Measurement, American Council in Education*, Washington, 1971.
- Vial, B. y R. Soto. “¿Predice la PAA el Rendimiento o el Éxito en la Universidad?”, *Revista de Administración y Economía*, P. Universidad Católica de Chile. 48 (2002). ☐

TÉCNICAS DE MEDICIÓN EN PRUEBAS DE ADMISIÓN A LAS UNIVERSIDADES*

Francisca Dussailant

Las pruebas de admisión a la educación superior en Chile han sido objeto de un intenso debate en los últimos doce meses. Entre otros aspectos, se ha sugerido aplicar en estas pruebas la teoría de respuesta al ítem (IRT, por sus siglas en inglés), en reemplazo de la teoría clásica de medición (TCM) que se usó en Chile en el marco de la PAA y que sigue utilizándose en Estados Unidos (SAT), en Suecia (SweSat), Israel (PET) y en el mundo entero en las pruebas de admisión universitaria.

En este trabajo se describen los elementos centrales de IRT, los modelos más utilizados dentro de esta teoría y algunas de sus aplicaciones, y se introduce el concepto de información de un ítem. A su vez, se presenta una recopilación de estudios e investigaciones que examinan en profundidad cuán robusta es IRT a las violaciones de los supuestos que la sustentan, cuán reales son las propiedades que la distinguen de la teoría clásica, y se analizan algunas de las conse-

FRANCISCA DUSSAILANT. Ingeniera Civil Industrial, P. Universidad Católica de Chile. Master en Educación, University of North Carolina-Chapel Hill. Investigadora del Centro de Estudios Públicos.

* Una versión preliminar de este trabajo apareció como *Documento de Trabajo* N° 347 (febrero de 2003), Centro de Estudios Públicos, Santiago.

La autora agradece a Carmen le Foulon y Harald Beyer por sus importantes aportes y sugerencias durante la preparación de este trabajo.

cuencias prácticas que traería su implementación en una prueba a gran escala, especialmente en el caso de que ésta tenga consecuencias para el examinado.

Del estudio realizado se concluye que IRT es una teoría promisorio, pero que *no* es la panacea que vendrá a solucionar todos los problemas que existen en medición educacional. Se sugiere, por lo tanto, que ella sea aplicada con cautela y como complemento a la teoría clásica, sobre todo en el caso de pruebas a gran escala con consecuencias para los estudiantes. Así lo recomendaría también la experiencia internacional, que muestra que se ha preferido mantener la teoría clásica como modelo dominante a partir del cual se calculan los puntajes en pruebas de admisión a la universidad, dejando IRT solamente para análisis secundarios.

1. INTRODUCCIÓN

Las teorías de medición sirven como marco teórico en el diseño e implementación de pruebas. Estas teorías entregan la metodología para la asignación de puntajes, proveen mecanismos para determinar las características de las preguntas o ítems, y a partir de ellas se derivan métodos para realizar otros análisis de interés. En este trabajo se analizarán dos de las principales teorías que se utilizan en el ámbito de medición educacional: la teoría clásica y la teoría de respuesta al ítem. El objetivo principal del artículo es analizar en particular la teoría de respuesta al ítem, sus posibles ventajas, aplicabilidad y limitaciones. Las siguientes líneas de esta introducción presentan de manera sucinta los principales análisis y conclusiones del trabajo.

Antes que nada, se hace necesario definir algunos términos. Distintas pruebas miden diferentes características de los examinados, por ejemplo: conocimiento matemático, razonamiento científico, capacidad de memorización, vocabulario, etc. En medición educacional, por “habilidad” se entiende la característica del examinado que va a ser medida a través de la prueba y, por consiguiente, ésta se utiliza para representar de manera genérica cualquiera de esas características. Por otro lado, en lo que respecta a los ítems de una prueba, hay parámetros (o descriptores) que permiten describir sus atributos particulares. Los descriptores de un ítem que suelen ser más frecuentemente utilizados y, por consiguiente, calculados o estimados son: a) nivel de “dificultad” (o sea cuán complicada es la pregunta) y b) nivel de “discriminación” de un ítem (o sea la capacidad que tiene éste para distinguir a los alumnos más aptos de los menos aptos).

1.1. Teoría Clásica

Una forma de medición que se utiliza con mucha frecuencia en pruebas a gran escala, como las de admisión a la universidad (por ejemplo, en el SAT de Estados Unidos y en la PAA en Chile), es la llamada teoría clásica (TC). En teoría clásica, el indicador de la *habilidad* de un estudiante corresponde al puntaje de éste en la prueba, construido a partir del número de respuestas correctas (o número de respuestas correctas netas) que obtuvo. Como indicador de la *dificultad* de una pregunta, en este sistema se utiliza la proporción de personas que la contestaron correctamente. Por otro lado, el índice de *discriminación* de un ítem se calcula como la correlación entre la respuesta a éste y el puntaje en la prueba total. Ambos descriptores de los ítems pueden calcularse ya sea en el contexto de una prueba piloto o experimental, o en la prueba definitiva u operacional.

Como se puede apreciar, en teoría clásica el grado de habilidad de una persona depende del grupo de ítems (vale decir, de su nivel de dificultad y discriminación) que contiene la prueba. Por ejemplo, si la prueba es fácil, un mismo alumno tendrá un puntaje mayor que si la prueba es difícil. Con esto resulta difícil hacer comparaciones entre estudiantes que han rendido pruebas diferentes. A su vez, los índices de dificultad y de discriminación de los ítems dependen del grupo de personas que rinden la prueba. Así, un mismo ítem puede ser catalogado como fácil si el grupo que rindió la prueba es excepcionalmente hábil, pero como difícil si el grupo que rindió la prueba es desaventajado. Con respecto a la discriminación, un ítem puede aparecer muy discriminatorio en el contexto de un grupo con nivel heterogéneo de habilidades, pero poco discriminatorio si el grupo que rindió la prueba es muy homogéneo (es decir, si todos los estudiantes tienen un nivel de habilidad similar). Esta dependencia de la habilidad de un estudiante con respecto al grupo de ítems de la prueba, junto con la dependencia de los descriptores de los ítems con respecto a las características del grupo, lo llamaremos *dependencia circular*.

Otra debilidad de la teoría clásica es que supone que la precisión con que se hace la medición es igual para todos los examinados, independientemente de su nivel de habilidad. Este supuesto es bastante discutible. Intuitivamente es claro que una prueba en que, por ejemplo, la mayor parte de sus preguntas son difíciles, va a distinguir más finamente entre dos personas con habilidad superior a la media que entre dos personas menos hábiles. Los que tienen habilidades inferiores obtendrán una estimación menos precisa de su habilidad, ya que son pocas o ninguna las preguntas de la prueba que responderán correctamente, y que, por lo tanto, servirían para

distinguirlos. Esta debilidad, junto con la dependencia circular que se genera en el cálculo de la habilidad de los examinados y la dificultad y discriminación de los ítems, han llevado a buscar un método que permitiese obtener una medida de la habilidad de los examinados que sea independiente de los ítems a que éstos se han enfrentado, una caracterización de los ítems independiente de la población a la que se aplican, y al mismo tiempo una medida más fiel de la precisión con que se está midiendo la habilidad. La satisfacción de estos requerimientos se ha intentado encontrar en la teoría de respuesta al ítem o IRT.

1.2. Teoría de respuesta al ítem (IRT, por sus siglas en inglés)

Los modelos IRT se centran en los ítems e intentan establecer, para cada uno de ellos, la *probabilidad* de ser respondidos correctamente. Esta probabilidad depende de la habilidad del examinado y de ciertas características de los ítems, entre las que pueden contarse su grado de dificultad y discriminación, y la probabilidad de ser respondido correctamente, por azar, por un individuo de muy baja habilidad. Hay varios modelos IRT de distinta complejidad. El modelo más simple es aquel que sólo diferencia los ítems según su grado de dificultad. Sin embargo, otros modelos permiten además incluir otros descriptores (o parámetros del ítem), como su grado de discriminación y la probabilidad de responderlo correctamente al azar.

Aparte de lo anterior, IRT hace posible conocer el nivel de “certeza” o “precisión” que un ítem aporta a la estimación para cada nivel de habilidad. En términos técnicos esto se llama “información” de un ítem. *Mientras mayor es la información que aporta el ítem a un determinado nivel de habilidad, mayor es la precisión en la estimación de ese nivel de habilidad.* Esto permite construir pruebas “a la medida” del objetivo educacional que se persigue.

La principal ventaja teórica de IRT es que mediante su utilización se lograría que un estudiante obtuviese siempre la misma estimación de su habilidad, independientemente de las preguntas (del banco de preguntas testeadas) que le tocó responder¹. También, con IRT un ítem tendría siempre los mismos parámetros que lo describen (dificultad, discriminación, etc.), independientemente del grupo que rindió la prueba. Esta notable propiedad se llama *invarianza* y es la piedra angular de la teoría de respuesta al ítem. A su vez, es la principal ventaja que distingue a IRT de la teoría

¹ Lo único que podría variar es el error de medición.

clásica. Sin embargo, es preciso destacar que la invarianza se cumple siempre y cuando se satisfagan ciertos supuestos y requisitos que se enunciarán a continuación.

Supuestos de IRT. (1) *Unidimensionalidad*: este supuesto consiste en que en una prueba todos los ítems están midiendo una y sólo una característica de los examinados. Esta propiedad está íntimamente ligada al supuesto número (2) *independencia local*, que postula que, dado un nivel de habilidad, las respuestas a los ítems no pueden estar correlacionadas entre sí. En otras palabras, si hay correlación entre preguntas, ésta sólo se explica por habilidad.

Otros supuestos importantes son: (3) que todos los alumnos que rindan la prueba hayan tenido *experiencias educacionales similares*; (4) que la prueba *no haya sido apurada*, y que (5) no haya “*efectos de contexto*” *no controlados*. Los “efectos de contexto” se refieren a que algunas preguntas se comportan de modo diferente según la posición que tengan en la prueba. Estos efectos se controlan adjudicando a la pregunta la misma posición en el pretest que en el test operacional.

Requisito básico de IRT. Aparte de los supuestos mencionados, para que se cumpla la propiedad de invarianza, y para que no haya problemas con la estimación de habilidad de los estudiantes, es de gran importancia que las predicciones del modelo se ajusten a los datos reales. Es decir, el modelo tiene que ser capaz de predecir con la mayor exactitud posible el comportamiento de los estudiantes frente a las distintas preguntas.

Ventajas y desventajas de IRT. La teoría de respuesta de ítem presenta una serie de potenciales ventajas sobre la teoría clásica. La principal de ellas es la *invarianza* de los puntajes de la prueba y de las características de las preguntas. También surge, gracias a las curvas de información, herramienta exclusiva de IRT, la posibilidad de optimizar el proceso de selección de preguntas según el objetivo educacional que se persigue². A su vez, con la teoría de respuesta al ítem se hace posible implementar Pruebas Adaptativas de Computador (CAT). En estas pruebas, cada examinado rinde una prueba “a su medida”, maximizando así la precisión en la medición. Otra ventaja de IRT es que presenta métodos alternativos para realizar ciertos análisis secundarios, como la detección de sesgos³, y presenta un

² Con IRT se hace posible controlar el error para los niveles de habilidad que se quieren medir con mayor precisión, al seleccionar los ítems más idóneos para este objetivo.

³ “Sesgos” corresponden a diferencias de desempeño para grupos de igual habilidad. Por ejemplo, un ítem estaría sesgado si los hombres con un cierto nivel de habilidad lo contestan de manera diferente que las mujeres que tienen ese mismo nivel de habilidad. En ese caso, un factor distinto a la “habilidad” que pretende medir la prueba, estaría jugando un rol importante en el desempeño de los individuos.

método alternativo de “equating” (proceso por el cual dos pruebas se hacen comparables).

Por cierto, todas las ventajas anteriores se pierden cuando los supuestos y requisitos de IRT no se cumplen. Para que haya invarianza, piedra angular de IRT y principal ventaja de ésta sobre la teoría clásica, es fundamental que exista, como señalamos anteriormente, unidimensionalidad e independencia local y que las predicciones del modelo se ajusten bien a los datos reales. Sin embargo, en muchas ocasiones la naturaleza de las disciplinas mismas les impide someterse a tales restricciones. Es por ello que en la práctica los supuestos y requisitos de IRT se transgreden a menudo. Este incumplimiento de los supuestos lleva no sólo a perder la invarianza sino que afecta directamente a la estimación de habilidad e introduce errores en aplicaciones secundarias de la teoría.

Veamos a continuación algunos ejemplos de transgresiones típicas de los supuestos y requisitos de la teoría.

De partida, se ha verificado que ciertas disciplinas son claramente multidimensionales, lo que atenta contra el requisito de unidimensionalidad de las pruebas. Por ejemplo, hay estudios que han demostrado que algunas pruebas de ciencias presentan varias dimensiones relevantes. Un estudio del NELS (National Educational Longitudinal Study), prueba de evaluación de la educación en Estados Unidos, detectó que las preguntas de la prueba apuntaban a tres dimensiones diferentes, claramente identificables. Éstas se denominaron “razonamiento y conocimientos básicos”, “razonamiento científico cuantitativo” y “razonamiento espacial-mecánico” (Hamilton *et al.* 1997, pp. 181-200, y Nussbaum *et al.*, 1997, pp. 151-173). Es esperable que en otras disciplinas, como ciencias sociales, también haya multiplicidad de dimensiones. Son pocas las soluciones que existen para corregir este problema. Una posibilidad sería seleccionar preguntas que representen sólo a una dimensión, pero esto implicaría descartar preguntas potencialmente valiosas para los objetivos de la prueba. Otra posibilidad es la de crear subpuntuajes, uno por cada una de las dimensiones testeadas. Con esto se incurre en una significativa disminución de la precisión en la medición, ya que cada subtest constará de muy pocas preguntas.

También se dan violaciones al supuesto de independencia local, las que se producen cuando las preguntas de una prueba están organizadas en torno a un estímulo común (Kolen y Brennan, 1995). Ejemplos de este tipo de preguntas son las comprensiones de lectura, donde a veces varios ítems se refieren a un mismo pasaje extenso. Lo mismo sucede cuando varios ítems se refieren al mismo gráfico o al mismo diagrama. En esos casos, es muy probable que las preguntas estén correlacionadas entre sí, aun cuando

se haya controlado por habilidad. Prescindir de este tipo de preguntas (comprensiones de lectura de textos largos, análisis de gráficos, etc.) puede ser muy costoso. Por ejemplo, hay tareas que sólo se pueden medir en torno a un texto largo, como la capacidad para seleccionar la idea principal, la capacidad de síntesis, la capacidad de jerarquizar y la capacidad de ordenar una secuencia.

En otras palabras, *si se quiere imponer el cumplimiento estricto de los supuestos de la teoría de respuesta al ítem se restringiría el campo de lo preguntable*, obligando a sacrificar buenas preguntas y a crear otras con contenidos ad-hoc. Esto puede redundar en preguntas de poca calidad, es decir, se deterioraría la validez de contenido de la prueba. *Por otro lado, si no se cumplen los supuestos de la teoría, se pierden todas las ventajas que ésta presenta frente a la teoría clásica*, y más aún, se puede incurrir en errores en la estimación de los puntajes de los estudiantes.

Continuando con el tema de las violaciones a los supuestos de IRT, cabe referirse al supuesto de las experiencias educativas similares. Es muy probable que en pruebas de gran escala diferentes estudiantes hayan sido sometidos a experiencias educativas muy distintas. Por ejemplo, si la habilidad que se va a medir son “conocimientos en biología”, dos estudiantes que tienen la misma habilidad en ese aspecto pueden haber asistido a clases muy diferentes. El profesor de María le dio mucho énfasis a la dimensión humana de la biología: María debió estudiar el cuerpo humano en detalle, las enfermedades más importantes, etc. Andrea en cambio tenía un profesor que le dio mucho más énfasis a la botánica y a todo lo que tiene que ver con el reino vegetal. Supongamos que Andrea y María son ambas muy buenas alumnas y saben el detalle de lo que se les ha enseñado. Ambas son igualmente “hábles” en biología. Sin embargo, lo más seguro es que una pregunta sobre polinización, por ejemplo, le va a parecer más fácil a Andrea que a María, y una pregunta sobre poliomielitis le parecerá más fácil a María que a Andrea. Lo anterior indicaría que, al momento de establecer la dificultad de una pregunta, va a ser importante en qué grupo de estudiantes se está testeando este nivel de dificultad. Si el grupo piloto tiene muchas Andreas la pregunta tendrá una dificultad diferente que si el grupo piloto tiene muchas Marías. Con ello, la dificultad de la pregunta deja de ser independiente de la muestra sobre la que se mide. En el fondo, IRT funciona en el supuesto de experiencias educacionales similares. Si no, la invarianza (o independencia de los descriptores de los ítems a la muestra) deja de ser real⁴.

⁴ Ver, por ejemplo, Miller y Linn (1988) citados en Fan (1998); Masters (1988) y Traub (1983) citados en Linn (1990); y Yen, Green y Burket (1987) citados en Green, Yen y Burket (1989), pp. 297-312.

Aun si se han cumplido todos los supuestos del modelo, es posible que el ajuste de éste a los datos no sea óptimo. Es decir, es posible que el modelo no sea capaz de predecir el comportamiento real de los estudiantes. Un pobre ajuste implica falta de invarianza, con lo que se pierde la principal ventaja de IRT por sobre la teoría clásica. Otro posible efecto de un modelo que no se ajuste a los datos es una estimación errónea de la habilidad de los estudiantes.

Otro problema de IRT tiene relación con la existencia de efectos de contexto. La dificultad y discriminación de una pregunta puede depender de la posición que ésta tiene en la prueba. Ha sido documentado que cierto tipo de preguntas presentan un efecto de “fatiga”, es decir, se hacen más difíciles mientras más avanzada es la posición que ocupan. Otras preguntas han demostrado tener un efecto de “práctica”, es decir, se hacen más fáciles si ha habido preguntas similares con anterioridad en la misma prueba. Estos efectos deben controlarse manteniendo siempre la posición de los ítems constante, tanto en las pruebas experimentales como en la prueba definitiva.

En el caso de pruebas de admisión universitaria, IRT trae consigo una serie de dificultades prácticas. Por una parte, la comprensión de los puntajes por parte de los estudiantes se dificulta. Con IRT puede suceder que estudiantes que tienen el mismo número de respuestas correctas, incorrectas y omitidas tengan puntajes diferentes, lo que suscitaría recelo e incompreensión por parte de los afectados.

Otro problema práctico tiene relación con la preparación de las pruebas por parte de los liceos. Al usar IRT en una prueba de admisión universitaria, se dificulta la creación de pruebas de ensayo, ya que los establecimientos educacionales no tienen la tecnología para estimar la dificultad y discriminación de las preguntas, ni los puntajes de los estudiantes.

En general, de implementarse IRT, la opinión pública se verá enfrentada a dificultades para verificar si el proceso de asignación de puntajes y de calibración de preguntas ha sido realizado de manera óptima.

El presente trabajo pretende introducir al lector en el tema de las teorías de medición, y en particular en la teoría de respuesta al ítem (IRT). La segunda sección describe de manera somera los fundamentos de la teoría clásica. La tercera sección trata en mayor detalle la teoría de respuesta al ítem, sus modelos y conceptos. En esta sección se estudiarán los distintos modelos de respuesta al ítem, se examinarán los supuestos y requisitos que necesita la teoría para funcionar de manera adecuada, se explicará cómo se estiman los parámetros del examinado (habilidad) y de las preguntas (dificultad y discriminación, entre otros), y se introducirá al lector en las funciones de información, herramienta de IRT que permite adecuar la selección de las preguntas de la prueba al objetivo educacional específico que ésta persi-

gue. La cuarta sección presenta un recuento y discusión sobre las consecuencias de las violaciones de los supuestos en IRT (unidimensionalidad, independencia lineal, entre otros). En esta sección también se examina la evidencia empírica que existe sobre la propiedad de invarianza. El objetivo de este examen es responder, al menos en parte, la interrogante de si en condiciones de medición reales se da realmente esta invarianza, tanto en las estimaciones de la habilidad del examinado como en la estimación de la dificultad y discriminación de los ítems. La cuarta sección, además, toca el tema de la calidad de las preguntas y la comprensión de los puntajes por parte de la población. Finalmente, la última sección presenta las conclusiones de este trabajo.

2. TEORÍA CLÁSICA DE MEDICIÓN

Esta teoría ha servido a la comunidad durante la mayor parte del siglo XX. Su mayor ventaja es que no depende de supuestos teóricos importantes y es un modelo relativamente simple. En esta teoría, cuando las preguntas de la prueba son dicotómicas⁵, la medida de habilidad del estudiante corresponde simplemente al número de respuestas correctas que obtuvo en la prueba (o en algunos casos a este número menos una fracción de las respuestas incorrectas)⁶.

La teoría clásica de medición presenta una alta dependencia de la muestra. Las estadísticas que describen los ítems de la prueba (p. ej., dificultad y discriminación) dependen del grupo de estudiantes que la rindieron. Como medida del nivel de dificultad de un ítem se utiliza, en el caso de preguntas con puntaje dicotómico, el cociente entre el número de estudiantes en la muestra que responden correctamente el ítem y el número total de estudiantes (es decir, la proporción de respuestas correctas). La discriminación de un ítem se expresa estadísticamente como la correlación entre los puntajes en el ítem y los puntajes totales de la prueba. Así, un ítem tendrá una alta discriminación cuando los estudiantes que lo responden correctamente son aquellos que tienen los puntajes más altos en la prueba y los que lo responden incorrectamente son aquellos que obtienen los menores puntajes.

⁵ Es decir, preguntas en que los puntajes son 0 ó 1, sin graduaciones intermedias.

⁶ Esta teoría también permite preguntas en las que se entregue puntaje parcial por una respuesta que no está completa, y permite que las distintas preguntas tengan un peso diferente en el puntaje total, peso determinado a priori según algún criterio definido. Sin embargo, para mayor simpleza, en este documento solamente se analizará la aplicación de la teoría clásica a preguntas dicotómicas con el mismo peso en el puntaje final.

La teoría clásica de medición también presenta una alta dependencia de los ítems, es decir, los puntajes que describen el desempeño de los alumnos dependen del grupo de ítems que se les entregó.

Quizá la principal dificultad que presenta la teoría clásica es que las características de los individuos no pueden separarse de las características de la prueba: ambas sólo pueden ser interpretadas en el contexto de la otra. La característica de los individuos que nos interesa la llamaremos la “habilidad” medida por la prueba, que en teoría clásica está definida por el “valor esperado del desempeño observado en la prueba de interés” (Hambleton, Swaminathan y Rogers, 1991). Como podemos ver, la habilidad de un examinado está definida solamente en términos de una prueba en particular. Cuando la prueba es difícil, el examinado aparecerá con una menor habilidad que cuando ésta es fácil. La dificultad de un ítem está definida como “la proporción de los examinados en el grupo de interés que contestan el ítem incorrectamente”. Así, que el ítem sea difícil o fácil depende de la habilidad de los integrantes del grupo. La discriminación de los ítems y la confiabilidad de los puntajes de la prueba también están definidas a partir del grupo de examinados. Así, las características de la prueba y de los ítems cambian cuando el grupo de los examinados cambia, y las características de los examinados cambia cuando el grupo de ítems que conforman la prueba cambia. Lo anterior hace difícil (aunque se han creado métodos⁷) comparar examinados que toman diferentes pruebas, y comparar ítems cuyas características se han obtenido usando diferentes grupos de examinados.

Otro problema de la teoría clásica de medición es que supone que el error de medición es constante para todos los examinados. El anterior es un supuesto poco plausible, ya que, como se explicó en la sección introductoria, la precisión de los puntajes varía de manera desigual según la habilidad de los examinados. Por ejemplo, consideremos a un estudiante que obtiene un puntaje cero. Este puntaje dice que el estudiante tiene una habilidad baja pero no queda claro cuán baja. Por otro lado, cuando un examinado contesta algunos ítems correctamente y otros de manera incorrecta, el puntaje de la prueba provee información más precisa sobre lo que el examinado puede o no puede hacer. Las medidas de confiabilidad⁸ en teoría clásica dejan también mucho que desear. Los diversos coeficientes de confiabilidad que

⁷ Hay algunos métodos para hacer “equating” con teoría clásica. De hecho, en la PAA Verbal se ecualizan las diferentes formas de ésta, y el SAT I utiliza métodos de “equating” por teoría clásica, complementándolos con IRT.

⁸ Medida de estabilidad de los puntajes que responde a la pregunta: ¿Se mantienen los puntajes cuando factores que no tienen relación con los propósitos de medición (contexto) varían?

una prueba entrega son un límite inferior para la confiabilidad o estimadores sesgados de ésta. Finalmente, la teoría clásica es una teoría orientada a la prueba y no al ítem. Así, con esta teoría se hace imposible predecir cómo se va a comportar en un ítem particular un examinado o un grupo de examinados. Preguntas como: ¿cuál es la probabilidad de que un determinado examinado conteste correctamente un cierto ítem?, son necesarias para el desarrollo de una serie de aplicaciones de pruebas. Este tipo de preguntas no se pueden responder desde la teoría clásica.

Por todo lo anterior se han buscado métodos alternativos de medición en los cuales las características de los ítems no sean dependientes del grupo examinado y los puntajes de los examinados no dependan del grupo de ítems utilizados en la prueba. Son deseables métodos que se expresen a nivel de los ítems y no a nivel de las pruebas, y que provean una medida de la precisión en la estimación de cada nivel de habilidad. IRT es uno de estos métodos, que ha experimentado un crecimiento exponencial en las décadas recientes.

3. IRT, MODELOS Y CONCEPTOS⁹

La teoría de respuesta al ítem descansa en dos postulados principales:

- El desempeño de un examinado en un ítem puede ser explicado por un grupo de factores llamados rasgos latentes o habilidades.
- La relación entre el desempeño de un examinado y el grupo de rasgos latentes se puede describir utilizando una función monotónicamente creciente llamada la “curva característica del ítem” (CCI). Esta función describe, para cada ítem, la probabilidad de responder correctamente según el nivel de habilidad del examinado. En general, es de esperar que mientras el nivel de habilidad aumenta, la probabilidad de responder correctamente a un ítem también aumente.

Hay muchos modelos posibles de respuesta al ítem, que se diferencian en la forma matemática de su CCI. Los modelos IRT contienen una serie de parámetros que describen un ítem y uno o más parámetros para describir al examinado. El primer paso en la aplicación de IRT es la estimación de estos parámetros. En este documento nos centraremos en los mode-

⁹ La estructura de este capítulo está basada principalmente en Hambleton, Swaminathan y Rogers (1991).

los más utilizados para describir ítems de respuesta dicotómica (es decir, que consideran sólo dos opciones: respuesta correcta o respuesta incorrecta). A pesar de que estos modelos no consideran la posibilidad de omitir una pregunta, hay una serie de posibilidades para incluir este tipo de respuestas en ellos¹⁰.

Los modelos matemáticos empleados por IRT se basan en una serie de supuestos sobre los datos en los que se aplica el modelo. El principal de estos supuestos es que sólo una habilidad o rasgo latente se mide por los ítems que conforman la prueba. El supuesto anterior se denomina *unidimensionalidad*¹¹. La *independencia local* es un requisito necesario pero no suficiente para que se cumpla la unidimensionalidad. Es por ello que muchos autores (p. ej. Hambleton *et al.* 1991) mencionan esta propiedad, que describiremos más adelante, como otro de los supuestos de IRT. Otro supuesto que se hace en todos los modelos IRT es que la CCI refleja la verdadera relación entre las variables no observables (habilidad) y las observables (respuestas a los ítems).

Se desprende del párrafo anterior que un requisito para que IRT funcione es que los supuestos enunciados se cumplan adecuadamente. Otro requisito fundamental tiene que ver con el ajuste del modelo a los datos. Un determinado modelo de respuesta al ítem puede ser o no adecuado para un conjunto particular de ítems. Si el patrón real de comportamiento de los estudiantes frente al ítem se diferencia significativamente de las predicciones que el modelo hace, entonces se puede afirmar que el ajuste no es el adecuado. En cualquier aplicación de IRT es esencial examinar el ajuste del modelo a los datos.

Cuando un determinado modelo de IRT se ajusta adecuadamente a éstos, y cuando no hay una violación importante del supuesto de unidimensionalidad, aparecen varias características deseables de IRT:

- Los estimadores de la habilidad de un examinado serán siempre iguales (excepto por errores de medición), independientemente del grupo que rinde la prueba, y los estimadores de los parámetros de los ítems obtenidos para diferentes grupos de examinados serán los mismos. La propiedad anterior, piedra angular de IRT, es la invarianza tanto de los parámetros de los ítems como del rasgo latente del examinado que se pretende medir;

¹⁰ El tratamiento de las respuestas omitidas está desarrollado, por ejemplo, en Lord (1980), pp. 225-231.

¹¹ Hay modelos multidimensionales para IRT, pero aún no están suficientemente desarrollados. En general, los modelos que se han aplicado en pruebas de gran escala son modelos unidimensionales.

- Con IRT los errores estándares de estimación¹² de habilidad son diferentes para cada nivel del rasgo latente.

A continuación se discutirán en detalle los supuestos de unidimensionalidad e independencia local.

Unidimensionalidad

Un supuesto común de IRT es que sólo una habilidad es medida por el grupo de ítems que conforman una prueba. Este supuesto no se cumple estrictamente en ningún caso debido a que varios aspectos cognitivos, de personalidad y operacionales siempre afectan de una u otra manera al desempeño en las pruebas. Entre estos aspectos se puede contar la ansiedad, nivel de motivación, rapidez, tendencia a responder al azar cuando no se sabe la respuesta, y otras destrezas cognitivas diferentes de la que supuestamente se está midiendo. Lo que de verdad importa es que entre todos los rasgos que afectan al desempeño en una prueba, uno de ellos sea claramente dominante sobre el resto. Existen modelos IRT multidimensionales, pero no están suficientemente desarrollados como para su implementación.

Independencia local

Este supuesto implica que no existen dependencias¹³ entre ítems que no sean atribuibles al rasgo latente que se está midiendo, lo que en términos estadísticos se traduce en que, para un individuo o un grupo de individuos de la misma habilidad, la probabilidad de que presenten un determinado patrón de respuestas en un grupo de ítems es igual al producto de las probabilidades de respuesta para cada ítem individual. Cuando se cumple el supuesto de unidimensionalidad, se obtiene también independencia local: en ese sentido los dos conceptos son equivalentes. Sin embargo, es posible que exista independencia local sin haber unidimensionalidad, cuando todos los rasgos latentes que influyen el desempeño son tomados en consideración. Por ejemplo, un ítem de una prueba de matemáticas que requiere un alto nivel de comprensión de lectura va a ser respondido incorrectamente por estudiantes que tienen poca habilidad lectora, sin importar sus destrezas

¹² En IRT el concepto “error estándar de estimación” de la habilidad es similar al de “error estándar de medición” de la teoría clásica.

¹³ En lenguaje estadístico, que —dado el nivel de habilidad— no exista correlación entre las respuestas a los diferentes ítems.

matemáticas. En ese caso, la independencia local no se cumpliría. Sin embargo, si todos los examinados tienen el nivel lector adecuado¹⁴, entonces solamente la destreza matemática tendrá efecto en el desempeño. La independencia local no se cumple cuando el ítem de una prueba contiene una pista sobre la alternativa correcta, o cuando provee información relevante sobre la respuesta a otra pregunta. La habilidad para detectar la pista es una habilidad diferente de la que está siendo examinada, por lo que no habrá independencia local.

3.1. Modelos más populares de IRT

La curva característica de un ítem (CCI) es una expresión matemática que relaciona la probabilidad de responder correctamente un ítem con la habilidad medida por la prueba y las características del ítem. Es posible concebir un número infinito de modelos de IRT, pero hoy en día sólo unos pocos se usan normalmente. En general los distintos modelos se diferencian por el número de parámetros que describen los ítems. Los tres modelos unidimensionales más populares son los modelos logísticos de uno, dos y tres parámetros que describiremos a continuación.

3.1.1. Modelo logístico de un parámetro

Es uno de los modelos de IRT más utilizados. Su CCI está dada por

$$P_i(\theta) = \frac{e^{(\theta - b_i)}}{1 + e^{(\theta - b_i)}} \quad i = 1, 2, \dots, n \quad [1]$$

donde

$P_i(\theta)$ es la probabilidad de que un examinado con habilidad θ conteste el ítem i correctamente.

b_i es el parámetro de dificultad del ítem i .

n es el número de ítems en la prueba.

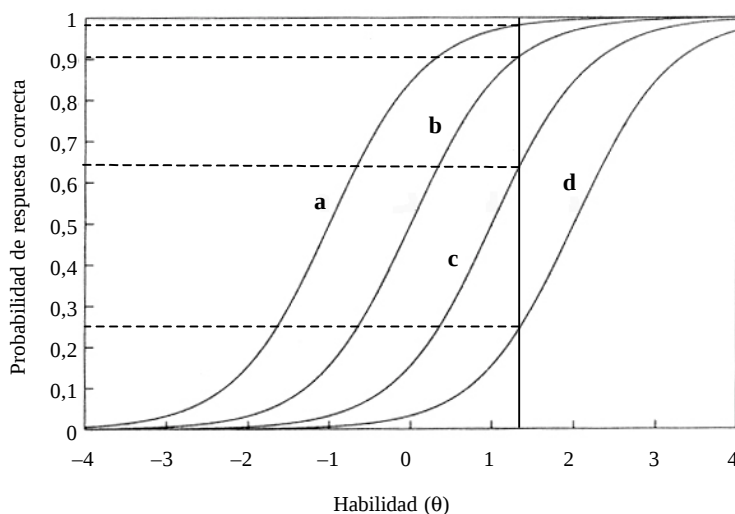
En este modelo, b_i corresponde al punto en la escala de habilidad en el cual la probabilidad de responder correctamente es 0,5. Cuanto mayor es el valor de b_i , tanto más difícil es el ítem, es decir, mayor la habilidad que se requiere para que un examinado tenga una probabilidad del 50% de responder correctamente.

¹⁴ En otras palabras, si controlamos por comprensión lectora.

Cuando los valores de habilidad¹⁵ de los estudiantes se transforman para tener una media de 0 y una desviación estándar de 1, los valores de b_i varían típicamente entre $-2,0$ y $2,0$. Ítems con b_i cercanos a $-2,0$ corresponden a ítems muy fáciles para el grupo de examinados a partir de los cuales se realizó la estandarización de la escala, e ítems con b_i cercano a $+2,0$ son ítems muy difíciles.

A continuación se presentan las curvas características para cuatro ítems de distinta dificultad que han sido modelados de acuerdo al modelo logístico de un parámetro. Como se puede observar en la Figura 1, las curvas de este modelo son muy similares una a otra; la única diferencia entre ellas está dada por su posición en el eje horizontal.

FIGURA 1: CURVAS CARACTERÍSTICAS DE ÍTEMS SEGÚN EL MODELO LOGÍSTICO DE UN PARÁMETRO¹⁶



En la Figura 1, la curva **a** corresponde a un ítem de dificultad $-1,0$, la curva **b** a un ítem de dificultad $0,0$, la curva **c** a un ítem de dificultad $1,0$ y la curva **d** a un ítem de dificultad $2,0$. Por lo tanto el ítem **d** es el más difícil, mientras que el **a** es el más fácil. La línea vertical corresponde a las personas con habilidad de aproximadamente $1,3$. Como podemos ver, una persona que pertenezca a ese nivel de habilidad contestará correctamente la

¹⁵ Los valores de habilidad θ están entre $+\infty$ y $-\infty$.

¹⁶ Figura adaptada de Hambleton, Swaminathan y Rogers (1991), p. 14.

pregunta **d** con una probabilidad inferior a 0,3. Sin embargo, esta persona responderá **c** correctamente con una probabilidad cercana a 0,65, y responderá **b** correctamente con una probabilidad superior a 0,9. La misma persona, casi con certeza, responderá **a** de manera correcta (la probabilidad en este caso se acerca a 1).

En este modelo logístico de un parámetro, también llamado a veces modelo Rasch¹⁷, se supone que la dificultad es la única característica del ítem que influencia en el desempeño del examinado. Este modelo no presenta ningún parámetro que corresponda al índice de discriminación de la teoría clásica. Lo anterior es equivalente a suponer que todos los ítems tienen el mismo grado de discriminación. Este modelo tampoco toma en cuenta que los examinados de baja habilidad van a tender a responder al azar, acertándole a veces a la alternativa correcta en el caso de pruebas con preguntas de selección múltiple.

Claramente el modelo de un parámetro se basa en supuestos bastante restrictivos, por lo que no siempre se ajusta de manera adecuada a los datos.

3.1.2. Modelo logístico de dos parámetros

La CCI para este modelo es menos restrictiva ya que relaja una de las suposiciones del modelo de un parámetro: la suposición de que todos los ítems tienen igual discriminación. En el modelo de dos parámetros aparece a_i , que correspondería a una medida de la discriminación del ítem. A continuación se presenta la forma matemática de la CCI según este modelo.

$$P_i(\theta) = \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \quad i = 1, 2, \dots, n \quad [2]$$

Como podemos ver, este modelo se parece mucho al modelo de un parámetro, excepto por la presencia de dos elementos nuevos. El factor D es una constante que se introduce para que la función logística se acerque lo más posible a una normal. Normalmente el valor de D utilizado es 1,7.¹⁸

Como ya se mencionó, el otro elemento adicional del modelo de dos parámetros es a_i , que corresponde a la discriminación del ítem. La discriminación es una medida de la capacidad que tiene el ítem de distinguir entre

¹⁷ Aunque en estricto rigor el modelo Rasch es diferente matemáticamente al modelo logístico de un parámetro, muchas veces se utiliza el nombre del primero para denominar al segundo.

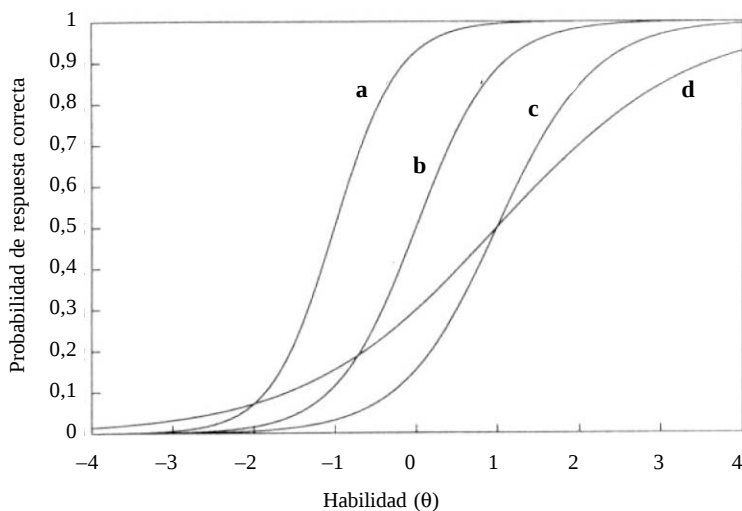
¹⁸ Se ha demostrado que cuando $D=1,7$, los valores de $P_i(\theta)$ para la ojiva normal de dos parámetros y el modelo logístico de dos parámetros difieren el valor absoluto en menos de 0,01 para todos los valores de θ . (Hambleton, Swaminathan y Rogers, 1991, p. 15).

un estudiante hábil y uno menos hábil. Si la probabilidad de contestar correctamente el ítem es similar para todos los niveles de habilidad, entonces estamos frente a un ítem poco discriminatorio. Si, al contrario, un estudiante con habilidad 1 tiene una probabilidad significativamente inferior de contestar correctamente el ítem que un estudiante con habilidad 2, entonces el ítem está discriminando adecuadamente a los estudiantes cuyas habilidades se ubican en esos rangos. El parámetro de discriminación a_i es proporcional a la pendiente de la CCI en el punto donde $\theta = b_i$. Los ítems con pendiente más pronunciada son más útiles para separar a los estudiantes en los diferentes niveles de habilidad. De hecho, la utilidad de un ítem para separar a los estudiantes con habilidades superiores a un cierto nivel θ de aquellos con habilidad inferior a θ , está dada por la pendiente de la CCI en θ .

Teóricamente, el parámetro de discriminación a_i está definido en la escala $(-\infty, +\infty)$. Los ítems con discriminación negativa (aquellos en que a menor nivel de habilidad del examinado, mayor la probabilidad de que conteste la pregunta correctamente) son eliminados automáticamente de la prueba. Es inusual que un ítem presente discriminaciones superiores a 2, por lo que el rango usual de estos parámetros es (0,2).

A continuación se presentan las curvas características de cuatro ítems que han sido modelados según el modelo de dos parámetros (Figura 2).

FIGURA 2: CURVAS CARACTERÍSTICAS DE ÍTEMS SEGÚN EL MODELO LOGÍSTICO DE DOS PARÁMETROS¹⁹



¹⁹ Figura adaptada de Hambleton, Swaminathan y Rogers (1991), p. 16.

Los valores de los parámetros para los ítems de la figura son los que siguen:

- Ítem **a**: Discriminación (a) = 1,5 Dificultad (b) = -1,0
- Ítem **b**: Discriminación (a) = 1,2 Dificultad (b) = 0,0
- Ítem **c**: Discriminación (a) = 1,0 Dificultad (b) = 1,0
- Ítem **d**: Discriminación (a) = 0,5 Dificultad (b) = 1,0

Como se puede apreciar en la Figura 2, el hecho de que las curvas presenten distintas pendientes indica que los valores del parámetro de discriminación a_i varían. Por ejemplo, podemos comparar dos ítems con dificultades idénticas pero diferente discriminación (ítems **c** y **d** en la figura) y notar que las curvas se comportan de manera significativamente diferente. La curva **c** es claramente más discriminante que la curva **d**, lo que se traduce en que en la primera es más fácil distinguir a los alumnos de habilidad inferior a 1,0 de aquellos con habilidad superior a 1,0. Como vemos, un estudiante de habilidad 0,5 responde el ítem **c** correctamente con una probabilidad cercana a 0,3, mientras un estudiante de habilidad 1,5 responde correctamente al mismo ítem con una probabilidad cercana a 0,7. Por otro lado, el estudiante de habilidad 0,5 responde correctamente **d** con una probabilidad de alrededor de 0,4, mientras el estudiante de habilidad 1,5 responde correctamente **d** con una probabilidad aproximada de 0,6. Con ello queda claro que la pregunta **c** distingue mejor entre estos dos estudiantes que la pregunta **d**, ya que, como vimos, la diferencia entre las probabilidades de contestar correctamente para estos dos estudiantes es de alrededor de 0,4 para la pregunta **c** mientras que para el ítem **d** esta diferencia se reduce a 0,2.

De la Figura 2 se desprende también que en este modelo de dos parámetros la probabilidad de contestar correctamente el ítem, para una persona de bajo nivel de habilidad es, en todos los casos, cero. Esto indica que el modelo de dos parámetros no considera la posibilidad, para un individuo de habilidad baja, de adivinar la respuesta correcta. Este supuesto puede ser plausible en el caso de preguntas de respuesta abierta pero no en el caso de preguntas de selección múltiple. Sólo si la prueba de selección múltiple es muy fácil es posible que un modelo de dos parámetros se ajuste adecuadamente (Hambleton, Swaminathan y Rogers, 1991), debido probablemente a que en ese caso los estudiantes de bajo nivel de habilidad no necesitarán adivinar, ya que conocerán la respuesta correcta.

3.1.3. Modelo logístico de tres parámetros

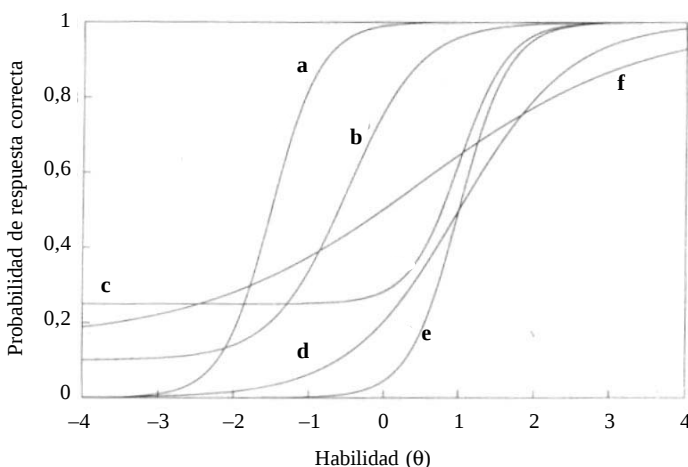
El modelo de tres parámetros incluye la posibilidad de que un estudiante de baja habilidad adivine la respuesta correcta. Con ello se relaja un supuesto del modelo de dos parámetros ($P(\theta) = 0$ para θ muy bajo), que es poco viable en el caso de preguntas de selección múltiple. La expresión matemática para este modelo es

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \quad i = 1, 2, \dots, n \quad [3]$$

El parámetro adicional que presenta este modelo y que lo diferencia del de dos parámetros es c_i . Este parámetro provee, en caso de ser necesario, una asíntota diferente de cero a la CCI, que corresponde a la probabilidad de que un estudiante con bajo nivel de habilidad adivine la alternativa correcta. Es común que c_i tome valores inferiores que el valor que resultaría si los estudiantes respondiesen al azar el ítem (0,2 en el caso de preguntas con cinco alternativas).

Es interesante notar que cuando hacemos $c_i = 0$, estamos volviendo al modelo de dos parámetros. Si adicionalmente fijamos la discriminación en $a_i = 1$, estamos volviendo al modelo de 1 parámetro.

FIGURA 3: CURVAS CARACTERÍSTICAS DE ÍTEMS SEGÚN EL MODELO LOGÍSTICO DE TRES PARÁMETROS²⁰



²⁰ Figura adaptada de Hambleton, Swaminathan y Rogers (1991), p. 18.

En la Figura 3, que presenta las curvas características de seis ítem, los valores de los parámetros para los ítems son los que siguen:

- Ítem a: Discriminación (a) = 1,8 Dificultad (b) = -1,5 Parámetro c = 0,0
- Ítem b: Discriminación (a) = 1,2 Dificultad (b) = -0,5 Parámetro c = 0,1
- Ítem c: Discriminación (a) = 1,8 Dificultad (b) = 1,0 Parámetro c = 0,25
- Ítem d: Discriminación (a) = 0,8 Dificultad (b) = 1,0 Parámetro c = 0,0
- Ítem e: Discriminación (a) = 1,8 Dificultad (b) = 1,0 Parámetro c = 0,0
- Ítem f: Discriminación (a) = 0,4 Dificultad (b) = 0,5 Parámetro c = 0,15

Como se puede apreciar, las curvas son bastante diferentes unas de otras según el valor de los parámetros que las definen.

3.2. Propiedad de invarianza

Como vimos anteriormente, la principal ventaja que diferencia a IRT de la teoría clásica es la independencia de la muestra, es decir, el que los parámetros de los ítems sean independientes del grupo de examinados para los cuales se estimaron, y los parámetros de habilidad sean independientes del grupo de ítems que contestó el individuo²¹. Es por ello que esta propiedad de “invarianza” es considerada en la literatura como la piedra angular de IRT (Hambleton *et al.*, 1991; Dorans 1990, entre otros).

Todas las ventajas de IRT sobre la teoría clásica se fundan en el supuesto de que la invarianza es una propiedad real y efectiva que en la práctica se cumple. El proceso de “equating”²² por IRT que incluso permite preecualizar²³ los parámetros de las preguntas, el proceso de detección de sesgos, el cálculo de errores de estimación de habilidad y el proceso de selección de preguntas idóneas para lograr un objetivo dado, son todas aplicaciones que con IRT se logran de mucho mejor manera que por teoría clásica, y que se fundamentan directamente en esta propiedad de invarianza de los modelos de respuesta al ítem.

En términos más técnicos, la invarianza significa que los parámetros que caracterizan a un ítem (a_i , b_i , c_i) no dependen de la distribución de

²¹ Excepto el error de estimación, que puede variar según los ítems que se eligen para estimar la habilidad.

²² Es decir, el proceso por el cual se logra que dos pruebas diferentes sean comparables.

²³ El pre-equating es un proceso de calibración de preguntas que permite que ellas estén calibradas en un banco de ítems incluso antes de ser utilizadas en una situación operacional.

habilidad del grupo de examinados con los que se estimaron, y el parámetro que caracteriza al examinado (θ) no depende del grupo de ítems utilizados en la prueba. Cuando el modelo IRT se ajusta a los datos, la invarianza implica que la misma CCI se obtiene independientemente de la distribución de habilidades de los estudiantes del grupo utilizado para estimar los parámetros de los ítems. Sin embargo, para la apropiada estimación de los parámetros del modelo es necesario que la muestra utilizada sea heterogénea. Lo anterior significa que en el grupo de individuos que conforman la muestra con la que se estimarán los parámetros debe haber personas de todas las habilidades, aunque no necesariamente distribuidas de la misma manera que la población objetivo.

Es importante hacer notar sin embargo ciertos puntos respecto de la propiedad de invarianza: esta propiedad rige solamente si el ajuste del modelo a los datos es perfecto. Incluso si el ajuste del modelo a los datos es perfecto, debido a que las CCI describen un comportamiento probabilístico, es muy improbable que se observe invarianza estricta²⁴.

Es por lo tanto de extrema importancia evaluar hasta qué punto se cumple la propiedad de invarianza, ya que cada aplicación de IRT capitaliza en esta propiedad de los modelos. Aunque la invarianza nunca podrá ser observada en el sentido estricto, es importante evaluar el grado en que se cumple esta propiedad al usar diferentes subgrupos de examinados. Por ejemplo, si dos muestras de distintas habilidades son extraídas de la población y los parámetros de los ítems son calculados para cada una de ellas, la congruencia entre los dos sets de estimadores es una medida del grado de cumplimiento de la propiedad de invarianza. Hasta la fecha, no existen criterios objetivos para la evaluación de esta propiedad, por lo que la decisión de si se cumple o no depende del criterio de quien aplica el modelo.

Es difícil sobrestimar la importancia de la invarianza de los parámetros en IRT. Como se dijo anteriormente, esta propiedad es la piedra angular de la teoría, y su cumplimiento hace posible aplicaciones tan importantes como el equating, creación de bancos de ítems, investigación de sesgo en los ítems, etc.

3.3. Estimación de los parámetros

El primer y más importante paso en la aplicación de IRT a los datos de una prueba es la estimación de los parámetros que caracterizan el modelo elegido. De hecho, la aplicación exitosa de la teoría de respuesta al ítem

²⁴ Es decir que en todas las instancias en las que se estimen los parámetros se llegue exactamente al mismo estimador.

depende de la disponibilidad de procedimientos satisfactorios para la estimación de los parámetros del modelo.

Tanto la habilidad de los examinados como los parámetros que describen a cada ítem son en principio desconocidos. Lo único que se conoce son las respuestas de los estudiantes a los distintos ítems. Por lo tanto, el problema de la estimación es determinar tanto el valor de θ para cada examinado, como el valor de los parámetros descriptores de cada ítem testeado. Si θ fuese una variable observable o conocida, la tarea de la estimación de los parámetros de los ítems sería bastante más sencilla de lo que de hecho es. El problema es que en IRT el parámetro de habilidad θ es un rasgo latente del examinado que no se conoce a priori.

La estimación de los parámetros se puede realizar de diversas maneras. Se busca encontrar el valor de los parámetros que produzca la curva de mejor ajuste a los datos. Esta estimación de los parámetros para lograr el mejor ajuste posible a los datos se realiza utilizando el criterio de máxima verosimilitud. A continuación se explicará a grandes rasgos en qué consiste este proceso de estimación.

Cuando los parámetros de los ítems son conocidos

En este caso lo que se pretende es encontrar el valor de habilidad (θ) que maximiza la probabilidad de tener un determinado patrón de respuestas²⁵. La función por maximizar es la función de probabilidad conjunta que explica el patrón de respuesta a los ítems que tuvo el estudiante. Debido a que rige el supuesto de independencia local, esta probabilidad conjunta no es más que la multiplicación de las probabilidades de respuesta de ese estudiante a cada uno de los ítems. Así, para un determinado estudiante, su habilidad θ se encuentra al maximizar

$$L(u_1, u_2, u_3, \dots, u_n | \theta) = \prod_{j=1}^n P_j^{u_j} Q_j^{1-u_j} = (P_1^{u_1} Q_1^{1-u_1}) \times (P_2^{u_2} Q_2^{1-u_2}) \times \dots \times (P_n^{u_n} Q_n^{1-u_n}) \quad [4]$$

Donde u_j corresponde a la respuesta que el estudiante dio a cada uno de los n ítems, y toma el valor de 1 ó 0 (pregunta correcta o incorrecta). P_j es la probabilidad de responder correctamente el ítem j para el estudiante que tiene habilidad θ , y $Q_j = 1 - P_j$ corresponde a la probabilidad de contestar incorrectamente²⁶. Nótese que esta función, llamada función de verosimilitud, se define sobre los n ítems respondidos por el examinado.

²⁵ El patrón de respuesta es la secuencia de respuestas que dio el alumno a las preguntas. Por ejemplo, el patrón {1,0,0,1,1,1,0,0,1,0} indica que el estudiante se enfrentó a diez ítems, respondiendo correctamente los #1, 4, 5, 6 y 9.

²⁶ Recordar que la forma de P_j está dada por la curva característica del ítem descrita en [1], [2] ó [3], según el modelo que se esté utilizando.

Por ejemplo, si tenemos un estudiante que rindió una prueba que constaba de tres ítems, y contestó correctamente el primero ($u_1 = 1$), incorrectamente el segundo ($u_2 = 0$) y correctamente el tercero ($u_3 = 1$), su *habilidad* se obtendría maximizando

$$P_1^1 Q_1^0 \times P_2^0 Q_2^1 \times P_3^1 Q_3^0 = P_1 \times Q_2 \times P_3 \quad [5]$$

Como los parámetros de las preguntas son conocidos, la única incógnita que tiene [5] es el nivel de habilidad θ del estudiante. P_i (y por lo tanto también Q_i) está definido en la CCI del modelo que se ha elegido para los ítems (ya sea [1], [2] ó [3]). Encontrar θ en este caso es directo, basta con encontrar el valor de la habilidad con la cual [5] es máximo.

Cuando los valores de θ son conocidos para cada uno de los estudiantes de la muestra, pero los parámetros de los ítems no se conocen

En este caso el proceso para encontrarlos es similar, pero no idéntico al descrito en el caso anterior. Se deben encontrar ahora los parámetros de los ítems que maximizan, para cada ítem, su función de verosimilitud. Esta función está dada por

$$L(u_1, u_2, u_3, \dots, u_K | \theta, a, b, c) = \prod_{i=1}^K P_i^{u_i} Q_i^{1-u_i} = (P_1^{u_1} Q_1^{1-u_1}) \times (P_2^{u_2} Q_2^{1-u_2}) \times \dots \times (P_K^{u_K} Q_K^{1-u_K}) \quad [6]$$

Donde u_i es la respuesta del estudiante i al ítem en cuestión. P_i es la probabilidad que el estudiante i tiene de responder correctamente el ítem, y $Q_i = 1 - P_i$ es la probabilidad de responder incorrectamente. Nótese que ahora la función de verosimilitud se define sobre los K estudiantes que respondieron el ítem.

En este caso, el ejemplo es también levemente diferente. Supongamos que tenemos tres estudiantes que respondieron un ítem. El primer estudiante lo respondió correctamente, el segundo también lo respondió correctamente y el tercero lo respondió incorrectamente. Si conocemos las habilidades de los tres estudiantes, debemos encontrar los parámetros (uno, dos o tres, dependiendo del modelo elegido) que maximizan

$$P_1^1 Q_1^0 \times P_2^1 Q_2^0 \times P_3^0 Q_3^1 = P_1 \times P_2 \times Q_3 \quad [7]$$

Como las habilidades de los estudiantes son conocidas, las incógnitas de [7] corresponden a los parámetros del ítem. P_i (y por lo tanto también

Q_i) está definido en la CCI del modelo que se ha elegido para el ítem (ya sea [1], [2] ó [3]). Para encontrar los parámetros, basta con encontrar el valor de ellos con el cual [7] es máximo.

Sin embargo, hay que considerar que en algún momento del proceso ni los parámetros que definen el ítem ni la habilidad de los examinados se van a conocer. Este problema se resuelve considerando simultáneamente todos los ítems y todos los examinados, es decir, maximizando la función de verosimilitud conjunta para ítems y examinados.

Cuando no se conocen los valores de θ ni los parámetros que describen los ítems

En este caso la estimación de ellos se complica. Ahora hay que encontrar los parámetros que maximizan la función de verosimilitud conjunta para todos los ítems y todos los examinados que conforman la muestra. Esta función por maximizar, cuando la independencia local se cumple, está dada por

$$L(u_1, u_2, u_3, \dots, u_K | \theta, a, b, c) = \prod_{i=1}^K \prod_{j=1}^n P_{ij}^{u_{ij}} Q_{ij}^{1-u_{ij}} \quad [8]$$

Como podemos ver, esta función considera el patrón de respuestas de K examinados a n ítems. El número de parámetros de habilidad que se busca es K (uno por cada examinado), y el número de parámetros de ítems es $3n$, $2n$ o n , dependiendo de si se pretende usar el modelo de tres, dos o un parámetro. Así, para el modelo de tres parámetros se deben encontrar los $3n+K$ parámetros que maximizan la ecuación [8].

Hay que notar que la estimación conjunta de los parámetros²⁷ genera un problema de indeterminación de éstos. Si en la CCI de, por ejemplo, el modelo logístico de tres parámetros se reemplaza θ por $\theta^* = \alpha\theta + \beta$, se reemplaza b por $b^* = \alpha b + \beta$ y a por $a^* = \alpha a + \beta$, la probabilidad de responder correctamente se mantiene sin cambiar, es decir, $P(\theta) = P(\theta^*)$. Como α y β son constantes arbitrarias, no existe un único conjunto de parámetros que maximice la función de verosimilitud [8]. Este problema puede ser eliminado eligiendo una escala arbitraria de habilidad o de dificultad (recordar que θ y b están en la misma escala). Por ejemplo se puede fijar la media y desviación estándar de los K valores de habilidad (o los n valores de dificultad), en 0 y 1 respectivamente. No debe olvidarse eso sí que se está eligiendo una escala arbitraria. Los parámetros de nuevas preguntas testeadas en una nueva muestra estarán por lo tanto en una nueva escala. Es

²⁷ Es decir, la estimación al mismo tiempo de los parámetros que describen a los ítems (a , b y c), y los parámetros que describen a las personas (θ).

por ello que éstos deben calibrarse de manera que sean comparables con los de la muestra original.

Una vez que se ha eliminado la indeterminación al fijar la escala, se calculan los valores de los parámetros que maximizan [8]. Para ello se puede utilizar el procedimiento de *estimación conjunta de máxima verosimilitud*, que corresponde a un proceso en dos etapas. En la primera etapa se eligen valores iniciales para los parámetros de habilidad. Usualmente se utiliza como valor inicial de θ el logaritmo de la razón entre número de respuestas correctas y el número de respuesta erradas del examinado. Estos valores son estandarizados para eliminar la indeterminación y con ellos se calculan los parámetros de los ítems.

En la segunda etapa, utilizando los parámetros de los ítems recién calculados, se estiman los parámetros de habilidad. Este proceso se repite reiterativamente hasta que los valores de los estimadores obtenidos en dos etapas consecutivas no cambien significativamente. El procedimiento de *estimación conjunta de máxima verosimilitud*, a primera vista atractivo, tiene una serie de desventajas. Primero, el método no puede estimar habilidad de un estudiante en el caso de tener todas las respuestas correctas, o todas incorrectas²⁸. Segundo, no hay estimadores para los parámetros de los ítems que han sido respondidos correctamente (o incorrectamente) por todos los examinados. Los ítems y los examinados que presenten estos patrones deben ser eliminados de la estimación. Tercero, para los modelos de dos y tres parámetros, el procedimiento no produce estimadores consistentes de los parámetros de habilidad y discriminación, aunque cuando la muestra es suficientemente grande este problema se reduce. Finalmente, en el caso del modelo de tres parámetros, el procedimiento numérico²⁹ para encontrar los estimadores puede fallar.

Por todo lo anterior, se han desarrollado métodos alternativos para la estimación de los parámetros. Uno de ellos es el de *estimación de máxima verosimilitud marginal*. Con este método, los parámetros de los ítems se estiman sin necesidad de hacer referencia a los parámetros de habilidad (θ), y los estimadores obtenidos son consistentes. Sin embargo, para utilizar este método se recurre a una suposición adicional: la muestra de estudiantes utilizada para hacer la estimación es representativa de la población de interés, y por lo tanto sus habilidades siguen una distribución conocida. Este método de estimación debe utilizarse con muestras suficientemente grandes de examinados.

²⁸ Algunos investigadores consideran ésta una seria desventaja del método. Otros no lo perciben como un tema problemático. En general, no hay consenso al respecto.

²⁹ Normalmente el procedimiento numérico utilizado en la estimación corresponde a la versión multivariable del método de Newton-Raphson.

En algunas ocasiones, incluso el método de *estimación de máxima verosimilitud marginal* puede fallar, es decir el proceso puede no converger a un resultado satisfactorio incluso después de muchas iteraciones. Esta falla sucede en especial en la estimación del parámetro c , en el modelo de tres parámetros. Malos estimadores de c a su vez resultan en la degradación de los estimadores de los otros parámetros de los ítems, y de la habilidad. Existe un proceso de estimación bayesiana que solucionaría este problema.

3.4. Evaluación del ajuste del modelo a los datos

La teoría de respuesta al ítem (IRT) tiene un gran potencial para resolver muchos de los problemas relativos a la medición mediante pruebas. Sin embargo, el éxito de las aplicaciones específicas de IRT no está asegurado por el solo hecho de calcular los parámetros de las preguntas según alguno de los modelos antes mencionados. Las ventajas de IRT se obtienen solamente cuando el ajuste de los datos al modelo es adecuado. Cuando los datos se ajustan pobremente al modelo, los parámetros de los ítems y los estimadores de habilidad de los examinados no tendrán la propiedad de la invarianza.

Hambleton, Swaminathan y Rogers (1991) recomiendan que los juicios sobre el ajuste del modelo a los datos deben estar basados en tres tipos de evidencia:

1. validez de los supuestos del modelo elegido;
2. nivel en el que se cumplen las propiedades esperadas del modelo (principalmente la invarianza de los parámetros);
3. precisión en las predicciones del modelo utilizando datos reales y, si es necesario, datos simulados;

La modelación de los datos con más de un modelo, para poder comparar resultados de ajuste, es una actividad especialmente útil para elegir el modelo apropiado. A continuación se detalla cada uno de los tres puntos anteriores.

3.4.1. Revisión de los supuestos

El principal supuesto para los modelos IRT aquí presentados es la unidimensionalidad. Otro supuesto importante es que la administración de la prueba no fue apurada (es decir, hubo el tiempo suficiente para contestar

las preguntas). Si la prueba fue apurada, entonces los parámetros de las preguntas dependerán no sólo de la habilidad latente del estudiante, sino de su rapidez. Además, cuando la prueba es apurada entra en juego fuertemente un factor de contexto: la dificultad estimada de una pregunta comienza a depender de la posición en la que está en una prueba, ya que, de encontrarse ésta al final, muchos individuos la omitirán o contestarán erróneamente, no por falta de conocimientos sino por falta de tiempo.

El modelo de dos parámetros tiene un supuesto adicional a los anteriormente enunciados: la probabilidad de que un estudiante de habilidad baja adivine la respuesta correcta es mínima o nula. En el modelo de un parámetro se supone además que los índices de discriminación son iguales para todos los ítems.

Hambleton *et al.* (1991) p. 57), presenta un recuento de métodos para examinar cada uno de estos supuestos. Entre ellos, se cita el análisis de factores como uno de los métodos que se utilizan normalmente en la evaluación de la unidimensionalidad.

3.4.2. Revisión de la invarianza

La invarianza se puede medir utilizando varios métodos directos. Por ejemplo, para revisar la invarianza del parámetro de habilidad (θ) se puede entregar a los examinados dos pruebas con niveles de dificultad marcadamente diferentes. Estimadores de la habilidad de los examinados se obtienen a partir de cada una de las pruebas. Si la invarianza se cumple, los estimadores de habilidad no debieran depender de la prueba a partir de la cual se calcularon.

Para medir la invarianza de los parámetros de los ítems se pueden calcular éstos a partir de dos grupos distintos de estudiantes. Si los estimadores son razonablemente similares, entonces la invarianza de los parámetros de los ítems se estaría cumpliendo.

3.4.3. Revisión de las predicciones

Hay variados métodos para chequear las predicciones del modelo. Uno de los más prometedores es el análisis de residuos. En este método se elige el modelo de respuesta al ítem, se calculan los parámetros y se realizan predicciones de desempeño para cada nivel de habilidad. Luego se comparan los valores predichos con los valores de los datos reales. La

diferencia entre el desempeño observado y el predicho es el “residuo bruto”. Este residuo se estandariza para tomar en cuenta el error de muestreo. El estudio de estos residuos puede entregar información valiosa. Por ejemplo, es importante que el tamaño de los residuos sea independiente del nivel de habilidad. Es decir, si los residuos en los niveles de habilidad superiores son grandes y los de niveles de habilidad inferiores son pequeños, eso puede ser indicio de que el ajuste no es el adecuado. Una manera de evaluar el ajuste a los datos es examinando la distribución de los residuos. Se considera que el modelo se ajusta adecuadamente cuando los residuos son pequeños y se distribuyen de manera similar que los residuos generados a partir de una simulación del modelo³⁰.

Pruebas estadísticas, generalmente chi-cuadrado, también se pueden aplicar para evaluar el ajuste del modelo a los datos. Para más detalle sobre estos métodos ver Hambleton, Swaminathan y Rogers, 1991, pp. 57-61.

3.5. La escala de habilidad (o del rasgo latente que se está midiendo)

El propósito fundamental de las pruebas es la asignación al examinado de un puntaje que refleje su nivel de logro de un objetivo medido por la prueba. Este puntaje debe ser interpretado con cuidado y debe ser válido en el contexto del uso que se le dará. En el marco de la teoría clásica, el puntaje asignado corresponde al número de respuestas correctas³¹. En teoría de respuesta al ítem, se supone que cada examinado tiene un nivel de habilidad θ que determina las probabilidades de responder correctamente un ítem. Así, basándose en el patrón de respuestas del examinado a la prueba, un puntaje de habilidad θ es asignado a cada examinado. Desafortunadamente³², la relación entre θ y el número total de respuestas correctas no es directa. Como fue descrito anteriormente, θ es independiente del grupo particular de ítems administrados y de la población a la que pertenece el examinado. Esta propiedad de invarianza es la que distingue a θ de los

³⁰ Los residuos generados a partir de la simulación de un modelo muestran cómo se distribuirían éstos si el modelo tuviese un ajuste perfecto. Por lo tanto, si los residuos reales se asemejan a los simulados, el ajuste del modelo es el adecuado.

³¹ En el caso de preguntas con puntaje dicotómico (0 ó 1). En algunos casos se agrega una penalización por cada respuesta incorrecta.

³² El problema que se genera al no existir una relación directa entre el número de respuestas correctas y el puntaje obtenido, es que se hace muy difícil para el estudiante entender el origen de su puntaje y muy difícil para los profesores simular una prueba de ensayo. Así, por ejemplo, es posible (y altamente probable) que dos individuos con el mismo número de respuestas correctas tengan diferentes puntajes finales en las pruebas.

puntajes asignados mediante teoría clásica. La escala de θ debe ser considerada como una escala absoluta con respecto al rasgo latente o habilidad que se está midiendo.

La escala θ original puede ser alterada mediante transformaciones lineales del parámetro que no alteran el cálculo de las probabilidades de responder correctamente los ítems (si es que los parámetros de los ítems son transformados también a su vez mediante el mismo proceso). Otra transformación de la escala de θ , y quizás una de las más importantes y utilizadas, es la *transformación a puntaje verdadero* (τ). En pocas palabras, el puntaje verdadero de un examinado con habilidad θ corresponde a la suma de las probabilidades de responder correctamente los ítems, evaluada en el valor de θ correspondiente. Si graficamos los *puntajes verdaderos* (τ) para todo el rango de posibles θ , lo que estamos haciendo es graficar la suma de las CCI utilizadas en las pruebas. La curva resultante se llama *curva característica de la prueba* (CCP). Mediante esta curva se mapea la relación existente entre la habilidad θ del examinado y su *puntaje verdadero* τ . El *puntaje verdadero* (τ) puede ser considerado como una relación no lineal de θ . Ya que $P_j(\theta)$ está entre 0 y 1, entonces el rango para el puntaje verdadero está entre 0 y n (con n = número total de ítems que conforman la prueba). La transformación de θ a *puntaje verdadero* (τ) tiene implicancias importantes. Primero que nada, se eliminan los puntajes negativos³³. La nueva escala, además, se mueve en el rango de 0 a n , que es más fácil de interpretar, aunque el puntaje τ obtenido por este método no corresponde a número de respuestas correctas, con lo que su interpretación se puede prestar a malos entendidos. La última implicancia de la transformación de θ a τ es que el τ de un examinado cuya habilidad θ es conocida puede ser computado para un set de ítems que nunca le han sido administrados. Esta característica es utilizada para las pruebas adaptadas al examinado³⁴ (hoy en día ampliamente utilizadas, por ejemplo, en Estados Unidos, en el GRE y el TOEFL, entre otros).

3.6. Funciones de información para los ítems y las pruebas

La teoría de respuesta al ítem (IRT) entrega una poderosa herramienta para describir los ítems y las pruebas, que mejora y facilita el proceso de selección de preguntas según los objetivos deseados. Éstas son las funcio-

³³ Recordar que θ se mueve en el rango $[-\infty, +\infty]$.

³⁴ Corresponde a pruebas en las que el computador entrega preguntas más fáciles o más difíciles, de acuerdo al desempeño que ha tenido el individuo en los ítems anteriores. De esa manera se hace posible estimar con menos preguntas y con mayor precisión la habilidad del examinado.

nes de información. Antes que nada, introduciremos un nuevo concepto ampliamente utilizado en IRT, el concepto de *información* de un ítem y de una prueba.

Cuando se habla normalmente de “tener información”, se evidencia que se tiene conocimiento de algo sobre un tema u objeto en particular. En estadística y psicometría, el término *información* tiene un significado similar, aunque un poco más técnico. En estadística, *información* se define como el recíproco³⁵ de la *precisión* con la que un parámetro es estimado (Baker, 2001). Así, al estimar un parámetro con precisión se obtendrá más conocimiento (y por ende información) de él que cuando el parámetro se estima de manera menos precisa. En IRT es posible determinar una curva de información para cada ítem, y una curva de información para cada prueba (que corresponde a la suma de las curvas de los ítems que la integran). La forma de las curvas de información de los ítems, $I_i(\theta)$, es

$$I_i(\theta) = \frac{\left[\frac{dP_i}{d\theta} \right]^2}{P_i(\theta)Q_i(\theta)} \quad i = 1, 2, \dots, n \quad [9]$$

donde $I_i(\theta)$ es la información que entrega el ítem i en el nivel de habilidad θ , $P_i(\theta)$ es la probabilidad de responder correctamente al ítem i dada por su CCI, y $Q_i = 1 - P_i$ es la probabilidad de responder de manera incorrecta el ítem i . En el caso del modelo de tres parámetros, la ecuación [9] puede simplificarse a

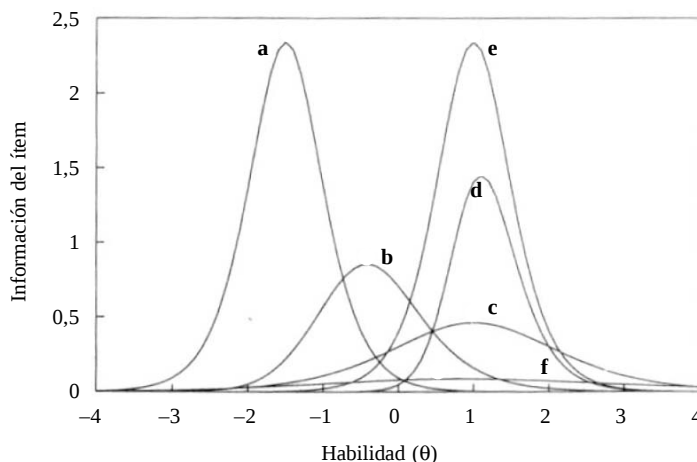
$$I_i(\theta) = \frac{2,89a_i^2(1 - c_i)}{[c_i + e^{1,7a_i(\theta - b_i)}][1 + e^{-1,7a_i(\theta - b_i)}]^2} \quad i = 1, 2, \dots, n \quad [10]$$

A continuación se presenta la Figura 4, que muestra las curvas de información correspondientes a los seis ítems que se presentaron hacia el final de la sección 3.1.

A partir de [10] y del análisis de la figura, se puede inferir el rol de los parámetros de los ítems en la función de información:

- un ítem entrega información máxima para $\theta = b$, es decir, cuando la dificultad del ítem se iguala a la habilidad del estudiante al que se está evaluando (esto es efectivo en los modelos de 1 y 2 parámetros,

³⁵ Por “recíproco” queremos decir que precisión e información están en relación inversa. La medida de “precisión” que se usa aquí es la desviación estándar de los estimadores. A medida que la “precisión” (desviación estándar) aumenta, menos precisa es la estimación.

FIGURA 4: EJEMPLOS DE CURVAS DE INFORMACIÓN DE ÍTEMS³⁶

en el modelo de tres parámetros, cuando $c > 0$, la máxima información se obtiene para θ levemente superior a b);

- la información de un ítem es generalmente alta cuando el valor del parámetro de discriminación (a) es alto. Esto se puede observar en el caso de los ítems **a** y **e** en la Figura 4, que corresponden a aquellos con mayores discriminaciones;
- la información de un ítem aumenta si c se acerca a 0 (esto se puede corroborar comparando las curvas de información para **c** y **e**, curvas que difieren solamente en el valor de c).

Como se aprecia en la Figura 4, un ítem con niveles bajos de discriminación es prácticamente inútil, ya que aporta muy poca información (observar el caso del ítem **f**). Sin embargo, es importante notar que incluso los ítems con niveles altos de discriminación entregan muy poca información en algunos rangos de habilidad³⁷. Podemos ver que **a**, por ejemplo, casi no entrega información sobre los rangos de habilidad superiores a 1, aunque entrega considerable información en el rango de habilidades que van entre -2,5 y -0,5 (recordar que la dificultad (b) del ítem **a** es -1,5, por lo que el

³⁶ Figura adaptada de Hambleton, Swaminathan y Rogers (1991), p. 93.

³⁷ Un ítem muy discriminante entregará mucha información para el rango de θ situado en los alrededores de $\theta=b$. Para rangos de θ muy diferentes al nivel de dificultad de la pregunta, ésta entregará información mínima o nula.

ítem estaría entregando más información en los alrededores de $\theta = b$, como es de esperar).

Las funciones de información juegan un rol importantísimo en el desarrollo de pruebas y evaluación de ítems, ya que muestran la contribución que los ítems hacen a la estimación de habilidad para los diferentes puntos del eje θ . Estas curvas permiten elegir el ítem que mejor se ajusta a los requerimientos de la prueba. Como se observó más arriba, un ítem tendrá poco valor si presenta un parámetro a bajo y un parámetro c alto, ya que en esos casos la información que entregaría sería mínima. La utilidad de un ítem específico va a depender de las necesidades específicas de la prueba que se está desarrollando. Un ítem puede proveer considerable información en un extremo del continuo de habilidad, pero si se requiere información de otro lugar de la escala, su valor se pierde. Sin embargo, se debe tener claro que si el modelo IRT elegido se ajusta pobremente a los datos, la curva de información puede inducir a conclusiones equivocadas.

Claramente, las funciones de información de los ítems proveen un nuevo enfoque para juzgar la utilidad de los ítems y construir las pruebas.

Función de información de una prueba

Esta función está dada por:

$$I(\theta) = \sum_{i=1}^n I_i(\theta) \quad i = 1, 2, \dots, n \quad [11]$$

Es decir, la información que entrega una prueba en un nivel de θ dado corresponde a la suma de las funciones de información de los ítems que conforman la prueba, evaluadas para ese valor de θ .

La cantidad de información provista por una prueba en un determinado θ está inversamente relacionada con la precisión con la cual la habilidad es estimada en ese punto.

$$SE(\hat{\theta}) = \frac{1}{\sqrt{I(\theta)}} \quad [12]^{38}$$

donde $SE(\hat{\theta})$ es el *error estándar de estimación*³⁹. Cuando se conoce la información de una prueba en un determinado θ , puede establecerse una

³⁸ La ecuación [12] es válida siempre que los parámetros del modelo IRT hayan sido calculados utilizando los métodos de máxima verosimilitud.

³⁹ El error estándar de estimación es similar a lo que en teoría clásica se denomina error estándar de medición.

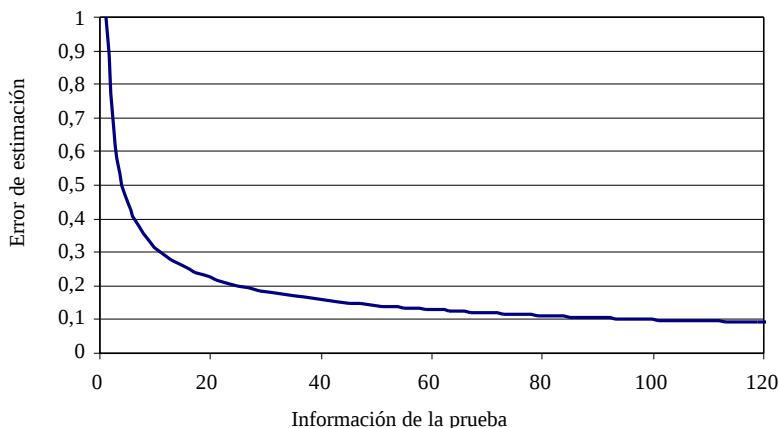
banda de confianza, dada por el error estándar de estimación, que ayuda a la interpretación del valor de θ encontrado. Es importante notar que el valor que toma $SE(\hat{\theta})$ varía con el nivel de habilidad, mientras que en la teoría clásica el error estándar de medición se considera constante.

La magnitud del error estándar depende en general de:

- El número de ítems en la prueba. A mayor el número de ítems, menor el error.
- La calidad de éstos. Mientras mayor es la información que entrega cada ítem, se considera que tiene mejor calidad, ya que aporta de mejor manera a disminuir el error de estimación de la prueba. (Mayor información se obtiene cuando el ítem presenta altas discriminaciones y valores bajos del parámetro c).
- La correspondencia entre la dificultad de los ítems y la habilidad del examinado. Pruebas demasiado fáciles o demasiado difíciles para el grupo que se va a examinar resultarán en errores estándares grandes, ya que el máximo de información que entrega la prueba estará en un rango de habilidad diferente al del grupo objetivo.

El valor de los errores estándares se estabiliza para niveles importantes de información. Así, aumentos de la información que entrega una prueba a, por ejemplo, más de 25, tiene sólo un leve impacto en la magnitud del error de estimación de θ . Lo anterior se puede comprobar en el siguiente gráfico que muestra cómo varía el error en distintos niveles de información (Figura 5).

FIGURA 5: INFORMACIÓN DE UNA PRUEBA VS. ERROR ESTÁNDAR DE ESTIMACIÓN DE LA HABILIDAD



Como muestra la Figura 5, el beneficio (en términos de disminución del error de estimación) que se obtiene al agregar una medida de información a una prueba es cada vez menor. En otras palabras, agregar el décimo ítem a una prueba de nueve ítems trae mayores beneficios que agregar el sexagésimo ítem a una prueba de 59 (suponiendo que todos los ítems son igualmente informativos). Sin embargo, no hay que olvidar que a pesar de lo anterior, con la prueba de 60 ítems se entregará una estimación de habilidad más precisa que con una prueba con menos preguntas. La última pregunta aporta cada vez menos, pero de todas maneras aporta algo a la precisión de la medición.

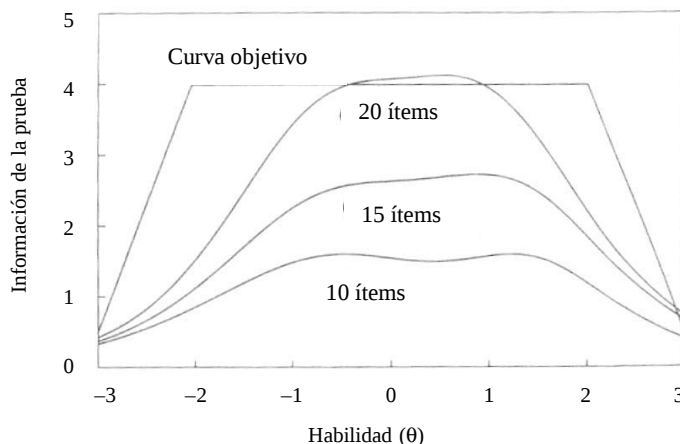
3.7. Construcción de pruebas

La teoría de respuesta al ítem ofrece un método de selección de ítems más poderoso que lo que ofrece la teoría clásica, principalmente debido a la propiedad de invarianza de los parámetros. Además el hecho de que la habilidad θ se mida en la misma escala que la dificultad b , hace posible seleccionar los ítems que son más útiles en ciertas regiones de la escala de habilidad. Quizá la principal ventaja de IRT es que permite seleccionar los ítems de acuerdo a la cantidad de información que aportan a la información total de la prueba, según las especificaciones de ésta. Ya que la información está relacionada con la precisión en la medición, es posible elegir los ítems que produzcan una prueba que mida con determinada precisión un cierto nivel de habilidad de interés.

El primer paso en la construcción de una prueba es determinar la forma deseada de su función de información. Para una prueba que apunta a un amplio rango de habilidades, la función de información objetivo debe ser relativamente plana, de manera de proveer información igualmente precisa en toda la escala de habilidad. Para una prueba en la que se establece un puntaje de corte (por ejemplo para separar expertos de novicios), la función de información objetivo debe tener un claro máximo cerca del puntaje de corte en la escala de habilidad.

La Figura 6 presenta la curva de información objetivo para una prueba en la que el interés es tener información sobre un amplio espectro de habilidades. En esta figura también se muestra el proceso a través del cual se intenta producir, con los ítems disponibles en el banco, una curva que se acerque lo mejor posible al objetivo.

FIGURA 6: CURVAS DE INFORMACIÓN PARA UNA PRUEBA CUYO OBJETIVO ES MEDIR UN AMPLIO ESPECTRO DE HABILIDADES⁴⁰

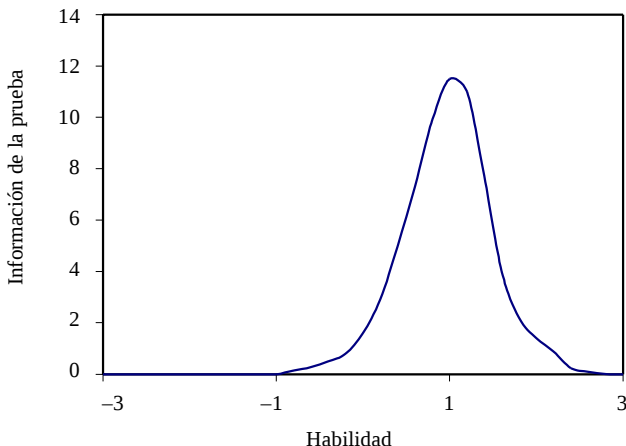


Para construir pruebas referidas a un criterio⁴¹, la forma de la curva de información objetivo será totalmente diferente, buscándose que el peak de información esté alrededor del nivel de habilidad de corte. Por ejemplo, en una prueba de razonamiento matemático queremos distinguir a aquellos que cumplen con un cierto mínimo de aquellos que no lo alcanzan. La habilidad que medirá esta prueba será “razonamiento matemático” y los estudiantes que superen un cierto nivel de corte serán aprobados y el resto será reprobado. Supongamos que el nivel de corte está en $\theta = 1,0$, entonces el objetivo de la prueba es distinguir con la mayor precisión posible a aquellos que superan ese criterio de aquellos que no lo hacen. No existe interés en distinguir, por ejemplo, a un estudiante de habilidad 2,0 de uno de habilidad 2,5, ni tampoco importa ser capaces de distinguir a uno de habilidad 0,0 de uno de habilidad $-1,0$. En este caso, la curva objetivo para la prueba será una curva cuyo máximo de información se ubicará en $\theta = 1,0$, y que entregará poca información sobre rangos de habilidad distantes del punto de corte. La Figura 7 ilustra una posible curva de información objetivo para una prueba como la recién descrita.

⁴⁰ Figura adaptada de Hambleton, Swaminathan y Rogers (1991), p. 104.

⁴¹ Como se explicó anteriormente, en estas pruebas se busca diferenciar a las personas que satisfacen un determinado criterio de aquellas que no. Para ello, se establece un punto de corte en el continuo de habilidad, alrededor del cual lo deseable es que exista un máximo de información.

FIGURA 7: POSIBLE CURVA DE INFORMACIÓN OBJETIVO PARA UNA PRUEBA REFERIDA A CRITERIO



A continuación presentamos las curvas de información para la prueba de admisión a las universidades en Suecia (SweSAT). Es importante dejar en claro que en el caso sueco, IRT no es utilizado para el desarrollo de las pruebas. La SweSAT todavía usa teoría clásica, aunque respalda a veces los análisis con IRT, por lo que fue posible conocer su curva de información.

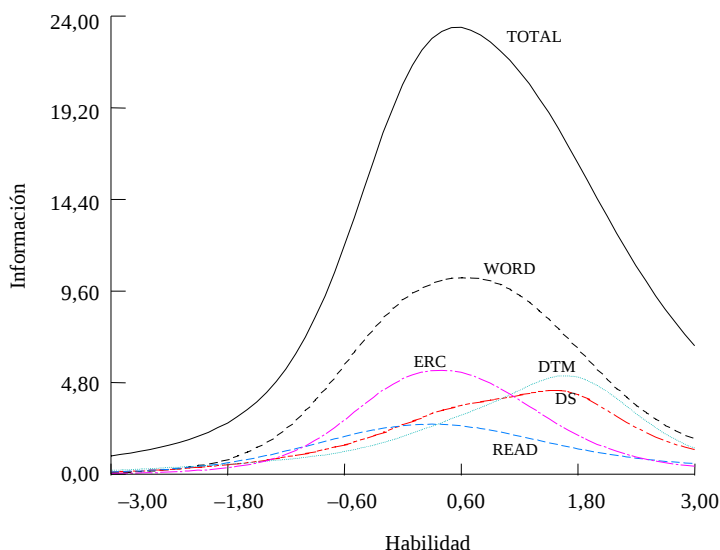
La SweSAT consta de cinco subtests:

- DS: Suficiencia de datos. Mide razonamiento matemático y consta de 22 ítems.
- DTM: Subtest de 20 ítems que miden la capacidad de interpretar diagramas, tablas y gráficos.
- ERC: Comprensión de lectura en inglés. 20 ítems.
- READ: Comprensión de lectura en sueco. 20 ítems.
- WORD: 40 ítems de vocabulario.

La Figura 8 presenta las curvas de información para los cinco subtests y para la prueba total.

Como podemos ver, el subtest WORD es el que más información estaría aportando. Lo anterior se debe a que éste corresponde al subtest con un mayor número de preguntas. Además, podemos ver que los peaks de

FIGURA 8: CURVAS DE INFORMACIÓN PARA EL SWESAT 2002



Fuente: Cristina Stage, Ph. D. Universidad Umeå, Departamento de Medición Educacional, Suecia.

información de todas las curvas están levemente desplazados hacia la derecha, en especial en el caso de los subtests DTM y DS. Este desplazamiento es comprensible, ya que lo que busca una prueba de admisión a las universidades es distinguir con precisión a aquellos que serán seleccionados para estudiar en ellas, que en general corresponden a estudiantes con niveles de habilidad superior a la media. Como podemos ver, en la curva de información TOTAL de la prueba, casi no hay información sobre estudiantes de habilidades inferiores a -1 , ya que ese rango de habilidades no interesa para los objetivos de la prueba. Sin embargo, la curva entrega considerable información en el rango de habilidades que va entre 0 y 2 , ya que es éste el rango que más interesa.

No existe una curva de información “óptima” en sentido general, ya que cada prueba persigue objetivos diferentes y por lo tanto las curvas de información que satisfacen mejor sus requerimientos son también distintas unas de otras. En el caso chileno, queda en evidencia la incompatibilidad entre una prueba que evalúe adecuadamente a la educación escolar, y una prueba que descubra de manera más precisa a aquellos estudiantes que

tendrían un buen desempeño en la universidad. Para establecer la curva de información óptima para un determinado test se deben definir los objetivos de éste y determinar, mediante una discusión en la que deben participar integrantes de todos los grupos involucrados⁴², el segmento en el continuo de habilidad sobre el que se necesita tener mayor información. El uso de funciones de información permite al encargado de desarrollar la prueba diseñar un test que se ajuste estrechamente a las especificaciones (suponiendo que se cuenta con un banco de ítems suficientemente grande y con ítems de calidad).

Sin embargo, existen algunos problemas con este enfoque (el de información de la prueba) que no pueden perderse de vista. Uno de ellos es que el uso de criterios estadísticos para seleccionar los ítems no asegura por sí solo la validez de contenido de la prueba. Es muy fácil sobrevalorar los criterios estadísticos y no tomar en cuenta el importante rol que juega el contenido de los ítems en el desarrollo de las pruebas. Más adelante se presenta una discusión detallada sobre este punto fundamental.

La estimación final y definitiva de los parámetros se hará en la prueba operacional, pero a esas alturas ya es tarde para seleccionar los ítems, por lo que las estimaciones hechas a priori en el estudio piloto servirán para este objetivo. Un problema de utilizar las curvas de información de los ítems en el desarrollo de pruebas es que valores altos de a (parámetro de discriminación) obtenidos a partir de grupos experimentales muchas veces son sobreestimaciones⁴³, resultando en curvas de información sesgadas (Hambleton, Swaminathan y Rogers, 1991). Es muy posible que una prueba construida con ítems de alto coeficiente a resulte, en la realidad, diferente de lo esperado. Una solución al problema de sobreestimación de los parámetros de discriminación es agregar algunos ítems extras a la prueba para que sirvan de compensación y la prueba finalmente entregue una información adecuada a los objetivos iniciales. Otra solución es utilizar en el estudio piloto muestras de estudiantes lo suficientemente numerosas en la estimación de los parámetros, de manera de obtener estimadores de éstos lo más correctos posible.

⁴² En el caso de una prueba de admisión a las universidades, entre los grupos interesados estarían las universidades que utilizarían los resultados de la prueba como instrumento de selección.

⁴³ En general, la propiedad de invarianza no se cumple estrictamente y muchas veces los parámetros, especialmente b y c y en menor grado a , sí varían según el grupo a partir del cual son estimados.

4. DISCUSIÓN Y ALGUNOS AVANCES DE LA INVESTIGACIÓN ACADÉMICA EN IRT

En la sección anterior se expuso a grandes rasgos lo que es la teoría de respuesta al ítem (IRT). A partir de lo expuesto surgen una serie de interrogantes que vale la pena estudiar más a fondo. Entre ellas está todo lo que tiene que ver con el cumplimiento de los supuestos sobre los que se funda la teoría, especialmente la unidimensionalidad y las implicancias que tiene en el modelo cualquier desviación de ella. Otro tema por profundizar es el que tiene que ver con la invarianza de los parámetros y hasta qué punto ésta se cumple en la práctica. También es de interés la discusión sobre el impacto práctico que puede tener la implementación de puntajes IRT en la población.

4.1. Unidimensionalidad (e independencia local)

Varios estudios académicos se refieren a éste el principal supuesto sobre el que se construye IRT⁴⁴. Al respecto, Childs y Oppler (2000, pp. 939-955), demuestran que el tema de la unidimensionalidad es un asunto de grados. Pequeñas desviaciones del supuesto tienen un impacto ínfimo en la aplicación del modelo. Los estudios de Childs y Oppler se basaron en un grupo de ítems que presentaban indicios evidentes de desviaciones del supuesto. Al comparar parámetros obtenidos en diferentes administraciones de los ítems (a diferentes grupos de personas), observaron que los parámetros de dificultad se mantuvieron estables, aunque los de discriminación no tanto, especialmente cuando se ajustó el modelo de tres parámetros. Los parámetros c fueron los que presentaron menor estabilidad. También observaron que los parámetros no variaban significativamente según si la calibración se hacía a partir de submuestras unidimensionales de ítems, o si se hacía a partir del pool completo de preguntas. Finalmente, los investigadores estudiaron la ordenación de los estudiantes al calcular un puntaje único a partir de los datos multidimensionales, y a partir de una combinación de subpuntajes calculados para los subgrupos unidimensionales de ítems. Las ordenaciones obtenidas se correlacionaron entre sí con valores de ρ superiores a 0,99. Childs y Oppler concluyen por tanto que pequeñas desviaciones de la unidimensionalidad no tienen un impacto significativo en la orde-

⁴⁴ Debe quedar claro que el hecho de que las desviaciones de la unidimensionalidad no sean detectadas al usar teoría clásica no significa que ésta sea una teoría superior en ese sentido: en teoría clásica también rige el supuesto de unidimensionalidad, pero en el caso de violaciones de ésta, las consecuencias prácticas son mucho menores que en IRT.

nación de los estudiantes que rinden la prueba. Green, Yen y Burket (1989) también sostienen que no es necesario que la unidimensionalidad se cumpla estrictamente para poder sacar provecho de los beneficios que traen consigo los modelos IRT. Otros estudios, que corroboran lo encontrado por Childs y Oppler tienden a demostrar que IRT es robusta a pequeñas desviaciones de la unidimensionalidad (Drasgow y Parsons, 1983; Harrison, 1986; Reckase, 1979)⁴⁵. Sin embargo, Ansley y Forsyth (1985), Doody (1987) y Reckase (1987)⁴⁶ muestran que la violación de la unidimensionalidad afecta seriamente a la estimación de los parámetros. Esto se complica más aun con la controversia sobre cuán apropiado es el análisis de factores⁴⁷ para analizar la dimensionalidad de IRT. No está claro si las dimensiones definidas por este método corresponden efectivamente a aquellas definidas por los rasgos latentes (Ansley, 1984)⁴⁸.

Krisci, Hsu y Yu (2001) examinan cuán robustos son los modelos de IRT a violaciones de la unidimensionalidad y concluyen que en el caso de aplicar modelos IRT a datos potencialmente multidimensionales se debe establecer cuál es la estructura de esa multidimensionalidad. Si hay una dimensión dominante y varias dimensiones menores, posiblemente IRT podrá ser aplicado. Se sugiere una varianza explicada de por lo menos 20% para el factor dominante. Cuando se dan varias dimensiones con predominancias similares, se debe evaluar la correlación entre ellas. Si las diferentes dimensiones están altamente correlacionadas ($> 0,4$) y las correlaciones son las mismas entre pares de dimensiones, es factible la aplicación de modelos IRT (Ackerman, 1989; Drasgow y Parsons, 1983; Harrison, 1986)⁴⁹. Si el análisis empírico de las preguntas lleva a la conclusión de que existe una fuerte multidimensionalidad y que las distintas habilidades medidas son más bien estadísticamente independientes, una posible solución es reconsiderar la definición del dominio de contenidos de la prueba, dividiéndolo en dos o más dominios relativamente amplios pero más homogéneos. Cada uno de estos subdominios debe tener un puntaje diferente. En el caso de uso de “subpuntajes”, se recomienda enfáticamente el uso de bandas de error tanto a nivel individual como a nivel grupal, de manera de evitar que exista una sobreinterpretación de las diferencias entre los diferentes subpuntajes debidas solamente al azar (Tate, 2002).

Del recuento anterior queda claro que IRT podría ser robusto a pequeñas desviaciones de la unidimensionalidad. Sin embargo, desviaciones

⁴⁵ Citados en Krisci, Hsu y Yu (2001).

⁴⁶ Citados en Krisci, Hsu y Yu (2001).

⁴⁷ Método ampliamente utilizado para evaluar la dimensionalidad.

⁴⁸ Citado en Krisci, Hsu y Yu (2001).

⁴⁹ Citados en Krisci, Hsu y Yu (2001).

“más que leves” del supuesto pueden afectar severamente a las principales aplicaciones⁵⁰ de IRT. Por ello es interesante estudiar cuáles son las fuentes más comunes de multidimensionalidad y de dependencia local (ambos supuestos, como se explica en la primera parte de este estudio, están altamente interrelacionados), para así facilitar la detección de estos potenciales problemas.

Kolen y Brennan (1995) explican que la independencia local implica que no hay dependencias entre los ítems que conforman una prueba que no sean atribuibles al rasgo latente o habilidad que el ítem está midiendo. Un ejemplo en el que esta característica probablemente no se cumpliría es cuando las pruebas están compuestas por conjuntos de preguntas basadas en un estímulo común (pasaje de lectura, gráfico, etc.). En este caso, el supuesto de independencia local sería violado porque los ítems asociados a un mismo estímulo probablemente estarán más interrelacionados entre ellos que con un ítem asociado a un estímulo diferente.

Lord (1980) considera que probablemente pruebas de ortografía, vocabulario, comprensión de lectura, razonamiento aritmético, analogías, series de números y pruebas que midan habilidades espaciales serán aproximadamente unidimensionales. Sin embargo previene que es muy fácil imaginar pruebas que violen este supuesto. Una prueba de conocimientos en química, por ejemplo, puede por una parte requerir entrenamiento matemático o destrezas aritméticas y por otra puede requerir manejo de conocimientos no matemáticos. Loyd (1988, pp. 135-143), declara que los ítems de las pruebas basadas en currículum probablemente se desviarán de la unidimensionalidad. Shavelson y Lau (2002) hablan del carácter multidimensional de las pruebas de ciencias. Al respecto Hamilton *et al.* (1997) y Nussbaum *et al.* (1997) encontraron evidencia de la estructura multidimensional del desempeño en ciencias, estableciendo tres dimensiones latentes a las que llamaron “conocimientos y razonamiento básicos”, “razonamiento científico cuantitativo” y “razonamiento espacial-mecánico”. Los ítems clasificados como de “razonamiento científico cuantitativo” correspondían a preguntas sobre contenidos de química, problemas que requerían computaciones matemáticas o ambos. Muchos incluían términos o procedimientos que se enseñan en cursos de química o física, como el cálculo de densidades o el balance de ecuaciones químicas. La mayoría de las preguntas de “razonamiento espacial-mecánico” incluían diagramas y requerían del uso de visualización del problema. Los ítems de “conocimientos y razonamiento básicos” preguntaban sobre conocimientos fácticos y razonamiento prin-

⁵⁰ “Equating”, construcción de pruebas, análisis de sesgos, construcción de bancos de ítems.

principalmente en las áreas de biología y astronomía. Lau *et al.* (2002)⁵¹ encontraron que variables motivacionales predecían el desempeño en ciencias incluso después de controlar por habilidad cognitiva y características demográficas. Kupermintz y Roeser (2002)⁵² encontraron evidencia similar a la de Lau *et al.* (2002). Loyd (1988) explica que, en un caso como el anterior, una manera de cumplir con los supuestos de IRT implicaría eliminar los ítems que estuviesen creando la multidimensionalidad. El problema potencial de esto es que la eliminación de ítems puede generar sesgos sistemáticos de contenido o de nivel de dificultad en las pruebas. Traub y Wolfe (1981)⁵³ son aun más enfáticos al sostener que existe una disparidad entre los objetivos de una prueba de contenidos y el concepto de unidimensionalidad. Advierten que los esfuerzos por lograr la unidimensionalidad seguramente van a restringir el rango de conocimientos examinado.

Otra fuente de violación de la unidimensionalidad son los efectos de contexto⁵⁴. Existe una extensa literatura al respecto, y las conclusiones a las que se han llegado son variadas. En general, se ha observado que en algunas situaciones de examinación (pruebas con límite de tiempo, por ejemplo) y en ciertos tipos de preguntas son más propensos a sufrir estos efectos. Leary y Dorans (1985, pp. 387-413), observan que en pruebas con límite de tiempo los estudiantes presentan un desempeño significativamente mejor cuando los ítems están ordenados de más fácil a más difícil. Yen (1980)⁵⁵ exploró los efectos del contexto en el cálculo de los parámetros de los ítems mediante IRT. Su estudio revela caídas consistentes en las correlaciones entre parámetros de discriminación calculados en contextos diferentes (en comparación con las correlaciones entre parámetros calculadas en contextos similares). También observó una baja en la estabilidad de los parámetros de dificultad. En particular, Yen encontró que los ítems de comprensión de lectura eran en general más difíciles cuando aparecían al final de una prueba que cuando aparecían al principio. Kingston y Dorans (1982)⁵⁶ realizaron un estudio similar y encontraron que algunas preguntas de la sección analítica del GRE presentaban “efecto de práctica” (es decir, su dificultad disminuía cuanto más cerca del fin de la prueba se ubicasen), mientras que los ítems de comprensión de lectura presentaban “efecto de

⁵¹ Citado en Shavelson y Lau (2002).

⁵² Citado en Shavelson y Lau (2002).

⁵³ Citado en Linn 1990.

⁵⁴ Los efectos de contexto tienen que ver con la ordenación de las preguntas en la prueba, y con el contexto en que se rinde la prueba (variaciones en el nivel de ansiedad, extensión total de la prueba, etc.).

⁵⁵ Citado en Dorans (1990).

⁵⁶ Citado en Dorans (1990).

fatiga” (es decir, la dificultad del ítem era mayor mientras más cerca del fin de la prueba se ubicase). Otros tipos de ítems estudiados por los investigadores no presentaron efectos significativos. Eignor y Cook (1983)⁵⁷ también observaron la falta de invarianza de los parámetros de los ítems de comprensión de lectura. Wightman (1981)⁵⁸ nota que, cuando la prueba tiene dos pasajes de comprensión de lectura, el nivel de desempeño en el segundo tiende a ser inferior al del primero. Una consecuencia potencialmente seria de la dependencia de la posición de los pasajes de comprensión de lectura es que la precisión real de un puntaje de prueba podría ser apreciablemente menor que la precisión aparente o nominal determinada por la teoría clásica o la teoría IRT. Como resultado de esto, una decisión aparentemente precisa sobre si un estudiante ha logrado un cierto objetivo puede de hecho estar equivocada debido a que el error de medición es mayor que lo calculado (Tate, 2002.)

Siguiendo con el tema de los efectos de contexto, Leary y Dorans (1985) notan que los ítems que miden aptitudes tienden a ser más sensibles a los cambios de orden que los ítems de pruebas de conocimientos. En una nueva investigación, Kingston y Dorans (1984) profundizan en el tema de estos efectos. En esta oportunidad descubren un leve efecto de práctica en los ítems de antónimos y un leve efecto de fatiga en los ítems de comprensión de lectura (GRE). En la sección analítica del examen GRE se observó que los ítems de análisis de explicaciones y los de diagramas lógicos presentaban efectos de práctica importantes y consistentes. Con lo anterior queda claro que los efectos de contexto, especialmente los que se refieren a la ubicación de las preguntas en la prueba, son específicos al tipo de ítem estudiado.

El diseño inapropiado de las condiciones de administración de la prueba puede resultar en la aparición de factores de multidimensionalidad. Por ejemplo, cuando el tiempo para contestar las preguntas es inadecuado para una proporción significativa de la población examinada, entra en juego el factor velocidad. Lo mismo sucede si la prueba es demasiado larga, en que entraría el factor “resistencia a la fatiga”. Incluso cuando el procedimiento de administración de la prueba es el adecuado para la prueba operacional, puede haber dudas respecto de los datos recogidos en las experiencias piloto utilizadas para evaluar empíricamente las pruebas. Por ejemplo, existe la preocupación de que la motivación de los examinados en una prueba experimental que “no cuenta” es menor que en la prueba operacional. Esta diferencia puede producir conclusiones inválidas en lo que

⁵⁷ Citado en Dorans (1990).

⁵⁸ Citado en Leary y Dorans (1985).

respecta a variadas interrogantes empíricas, incluyendo la de la dimensionalidad de la prueba⁵⁹ (Tate, 2002).

En el trasfondo de la discusión sobre unidimensionalidad está la siguiente cuestión: podría ser que una determinada prueba, muy apropiada desde el punto de vista pedagógico y de los objetivos educacionales que persigue, presente multidimensionalidad. En ese caso, es importante distinguir que los requerimientos para que IRT funcione con esa prueba no son siempre compatibles con los requerimientos reales de ésta. Por ejemplo, el hecho de que un pasaje largo de comprensión de lectura pueda ser fuente de multidimensionalidad no es razón suficiente para tomar la decisión de no hacer preguntas de este tipo en una prueba. El método estadístico a utilizar en el análisis de una prueba no debe primar por sobre consideraciones más importantes relacionadas con el contenido a medir y el cumplimiento de los objetivos reales que persigue la prueba en cuestión.

4.2. Invarianza

La propiedad de invarianza es la piedra angular de IRT, y lo que la distingue de la teoría clásica de medición. Es bastante sorprendente que estudios empíricos que examinen o comparen la invarianza de los parámetros en el contexto de ambos modelos sean tan escasos. Al respecto, Fan (1998, pp. 357-381), argumenta que “Al parecer la superioridad de IRT sobre la TCM (teoría clásica de medición) en este sentido ha sido dada por sentada por la comunidad de medición, y ningún tipo de escrutinio empírico ha sido considerado necesario. El silencio empírico en esta materia parece ser una anomalía”.

Stage (1998a) comparó los estadísticos de la sección de Comprensión de Lectura en inglés de la prueba SweSAT calculados en el marco de la teoría clásica y en el de IRT. Descubrió que los estadísticos que definen discriminación y dificultad tenían alta correlación entre los modelos. Stage también estudió la correspondencia entre los datos obtenidos en los pretests y los datos de la prueba operacional, no encontrando diferencias importan-

⁵⁹ Es posible que una prueba, en el contexto experimental (menos presionado), muestre, por ejemplo, dos dimensiones, pero que estas dos dimensiones se fundan en una en el contexto operacional. Por ejemplo, si la prueba es de ciencias, es muy posible que en la prueba piloto los estudiantes sepan o biología o física, según sus preferencias particulares, transformándose éstas en dimensiones diferentes. Sin embargo, en el contexto operacional, los estudiantes tienen más incentivo para rendir y estudiar la materia que no conocen. Por lo tanto, es posible que los que saben biología aprendan física y viceversa, transformándose entonces la medición en una medición unidimensional del rasgo latente “conocimiento científico”.

tes en lo que se refiere a dificultad de las preguntas⁶⁰, pero sí en la discriminación. En teoría clásica la correlación entre las discriminaciones calculadas en el pretest y las del test operacional fue de 0,57, mientras que en IRT fue de 0,34. La conclusión final de Stage fue que la supuesta superioridad de IRT sobre la teoría clásica no se verificaba con sus datos. La investigadora obtuvo resultados similares al estudiar la sección de Comprensión de Lectura en sueco del SweSAT (Stage, 1998b), y de la sección de Vocabulario (Stage 1999).

Miller y Linn (1988)⁶¹ reportaron también considerables diferencias en las curvas características de sus ítems, resultado de variaciones en la cobertura de instrucción de distintos grupos. Lo anterior sugiere falta de invarianza de sus parámetros. Masters (1988) y Traub (1983)⁶² sostienen que en pruebas de conocimientos los parámetros de los ítems van a ser sensibles a diferencias en la experiencia educacional. Otro estudio de Cook, Eignor y Taft (1988)⁶³ también reporta falta de invarianza en los estimadores de dificultad de los ítems, tanto en el contexto de la teoría clásica como en el contexto de IRT. Yen, Green y Burket (1987)⁶⁴ explican que algunos ítems pueden presentar parámetros marcadamente diferentes en ciertos establecimientos educacionales y distritos debido a patrones específicos de instrucción y aspectos curriculares. Por ejemplo, una prueba de aritmética puede ser esencialmente unidimensional entre grupos a los que se les ha enseñado prácticamente lo mismo. Sin embargo, si algún grupo de alumnos de tercer año ha pasado la unidad de multiplicación y otro grupo no, los estimadores de los parámetros para los ítems de multiplicación derivados de la muestra de estudiantes que conocen el tema van a ser sustancialmente diferentes de los parámetros derivados del grupo que no sabe multiplicar (Green, Yen y Burket, 1989).

Fan (1998) estudió la invarianza de los parámetros de dificultad y discriminación de los ítems utilizando para ello los datos del Texas Assessment of Academic Skills, que corresponden a pruebas basadas en currículum. Con respecto a los parámetros de dificultad, Fan encontró que los parámetros de teoría clásica eran levemente más invariantes que los de IRT. También observó que los parámetros de dificultad en ambas teorías eran más invariantes que los de discriminación, sin observarse una ventaja sistemática de un modelo por sobre otro. Cabe señalar que los resultados ante-

⁶⁰ Con teoría clásica, la correlación entre los estadísticos de dificultad del pretest y los del posttest fue de 0,87 y con IRT 0,88.

⁶¹ Citado en Fan (1998).

⁶² Citado en Linn (1990).

⁶³ Citado en Fan (1998).

⁶⁴ Citado en Green, Yen y Burket (1989).

rios no se deben a desviaciones de la unidimensionalidad ni a pobre ajuste entre modelo y datos. Ambos supuestos fueron chequeados por el autor y obtuvo resultados razonables de ajuste para los modelos de dos y tres parámetros, verificándose un nivel adecuado⁶⁵ de unidimensionalidad.

En general, existe evidencia mixta respecto de la consistencia de los parámetros de los ítems. Algunos investigadores han dicho que estos estimadores, al ser calculados a partir de diferentes muestras, han resultado razonablemente consistentes (Dorans y Kingston, 1985; Forsyth, Saisangjan y Gilmer 1981; Rentz y Barshaw, 1977)⁶⁶. Otros han encontrado que los estimadores varían considerablemente (Cook, Eignor y Taft, 1984; Loyd y Hover, 1980; Slinde y Linn, 1978)⁶⁷. Esta combinación de estudios sugiere que los modelos IRT no pueden ser aplicados sin un análisis cuidadoso de la situación. Cada situación de medición debe ser evaluada individualmente, para poder establecer la aplicabilidad del modelo (Loyd, 1988).

4.3. Impacto práctico de IRT

Las curvas características de los ítems y las curvas de información para las pruebas son herramientas valiosas y efectivas en la construcción de pruebas. Sin embargo, no debe nunca olvidarse que este tipo de herramienta no sirve de nada sin un buen banco de preguntas. Las herramientas estadísticas que IRT provee al análisis de los ítems son un complemento, pero no reemplazan ni permiten evadir el esfuerzo editorial inherente a la construcción de buenas preguntas (Green, Yen y Burket, 1989). Es importante que las preguntas no sólo exhiban los parámetros adecuados, sino que deben ser óptimas también en lo que se refiere a contenido y diseño, y ese tipo de evaluación no tiene nada que ver con las herramientas estadísticas aquí presentadas. Linn (1990) hace notar que en la descripción de los métodos de selección de ítems por IRT no se hace mención alguna del contenido de los ítems a ser seleccionados. No debe olvidarse que en pruebas de conocimientos las especificaciones de contenido son prioritarias. Al parecer, existe preocupación por parte de algunos académicos de que la sobrevaloración de la estadística lleve a distorsiones en las mediciones. Rudner (1998) enfatiza el punto: “los bancos de ítems y la teoría de respuesta al ítem no son la panacea que va a solucionar todos los problemas de medición. Sigue

⁶⁵ No está demasiado claro cuál es este “nivel adecuado”. La única sugerencia que se encontró fue la de Reckase (1979), citado en Krisci, Hsu y Yu (2001), que habla de que el factor dominante explica un 20% de varianza.

⁶⁶ Citados en Loyd (1988).

⁶⁷ Citados en Loyd (1988).

siendo importante el cuidado y esfuerzo en la creación de ítems. [...] Se debe hacer todo el esfuerzo posible para incluir sólo ítems de alta calidad en el banco de ítems. El mismo cuidado y esfuerzo debe ser dedicado a la creación de las preguntas. Los ítems obtenidos de fuentes externas deben ser evaluados cuidadosamente tanto en lo que se refiere a correspondencia con el currículum como a la calidad técnica de éstos.”

Otro problema práctico importante que IRT presenta a los profesionales de educación es la complejidad en la estimación de los parámetros y en general del modelo matemático en que se funda IRT. Al contrario de la teoría clásica, la aplicación de IRT a los ítems requiere del manejo de un modelo matemático complejo, y genera puntajes difíciles de interpretar. Para llevar a cabo el proceso se requiere de software computacional sofisticado y de personal calificado, lo que implica que los profesores y personas encargadas de preparar a los estudiantes para las pruebas probablemente no tendrán acceso a los parámetros de las preguntas que diseñen. Incluso si tuviesen estos parámetros, el proceso de estimación del puntaje del estudiante no es trivial, requiriéndose para ello un software especial. Esto dificulta sobremanera la posibilidad de hacer estimaciones a priori de cuál será el desempeño final de los estudiantes. Además, ni el estudiante, ni sus familiares, ni los profesores tendrán las herramientas necesarias para hacer un juicio independiente sobre cuán razonables les parecen los resultados en las pruebas. Todo esto añade un áurea misteriosa a los resultados de IRT (Lloyd, 1988).

Este misterio que rodea a IRT también se aplica a la escala de puntajes obtenidos por este método. La manera en que los parámetros que describen los ítems y el parámetro de habilidad θ del examinado están expresados no tiene ninguna relación directa con datos observables (es decir, número de respuestas correctas). No sólo es posible, sino altamente probable, que dos alumnos con exactamente el mismo número de respuestas correctas, incorrectas y omitidas tengan puntajes diferentes. Con ello, la generación de puntajes mediante IRT parece, para quien no maneja el modelo, un proceso mágico, lo que impide al ciudadano común hacer una revisión y chequeo de los resultados (Lloyd, 1988).

En el caso de Estados Unidos, donde existe una amplia gama de pruebas con consecuencias para admisión universitaria de pre y postgrado, la visualización de los problemas prácticos que trae IRT ha hecho que la principal compañía encargada del diseño de pruebas ETS⁶⁸ decida evitar en lo posible la asignación de puntajes y corrección de pruebas utilizando la

⁶⁸ Educational Testing Service. Encargados de diseñar, entre otras, las pruebas SAT I, SAT II, GRE, TOEFL, GMAT, PRAXIS, etc.

teoría de respuesta al ítem. Es por ello que *los puntajes de todas las pruebas de consecuencias administradas en formato de papel y lápiz son asignados por teoría clásica*. En el caso de pruebas adaptativas al computador, donde la teoría clásica no es una posibilidad, ETS y otras compañías privadas (por ejemplo Kaplan, Barrons, Petersons) han desarrollado software especializado para su preparación⁶⁹. Este software entrega ensayos computacionales de la prueba, de manera que el estudiante practique y obtenga estimaciones reales de su puntaje al final de cada práctica. Esto no sería posible mediante un facsímil impreso en papel, ya que estas pruebas utilizan IRT y por lo tanto la estimación de los puntajes no sería posible sin el software. En suma, los problemas prácticos que traen consigo los exámenes que utilizan IRT como principal teoría han llevado a que Estados Unidos haya hecho lo posible para mantener la teoría clásica y, en el caso de no ser esto posible, las instituciones encargadas del diseño de estas pruebas se han encargado de entregar a los estudiantes, muchas veces de manera gratuita, un software que les permita ensayar la prueba de manera adecuada.

5. CONCLUSIONES

Como hemos visto, IRT es una teoría que promete solucionar una serie de problemas que hasta ahora existían en la medición educacional. La principal ventaja que ofrece esta teoría es la invarianza de los parámetros que describen los ítems (dificultad, discriminación, probabilidad de responder correctamente al azar), y de los parámetros que describen a las personas (la habilidad θ en el caso de modelos unidimensionales).

La utilización de IRT en contextos de evaluación educacional es sin duda un aporte significativo que facilitará y perfeccionará la tarea de diseño e implementación de pruebas, especialmente aquellas de gran escala. IRT es una teoría emergente que está en continuo perfeccionamiento y es probable que en un futuro cercano las bondades de los modelos IRT sean aun mayores. Con todo, es importante poner una nota de cautela, ya que las bondades de la teoría pueden ser fácilmente sobredimensionadas.

De partida, no se debe olvidar que IRT no es necesariamente una alternativa a la teoría clásica de medición, sino que puede también ser visto como un complemento. Mientras más información se tenga de una prueba y

⁶⁹ En el caso de ETS, el Software POWERPREP se diseñó para preparar las versiones adaptativas al computador del GRE, GMAT y TOEFL. En los primeros dos casos POWERPREP es gratuito.

de sus ítems, mejor. Al usarse las dos teorías se están haciendo simultáneamente los análisis con ambas, con lo que, de fallar alguna, siempre estará la otra como respaldo.

Tampoco se debe olvidar que la teoría clásica tiene algunas ventajas significativas, especialmente en lo que se refiere a la comprensión de los puntajes que entrega. En general, sobre todo en el caso de pruebas con consecuencias, con estudiantes altamente motivados y familias altamente involucradas, la importancia de comprender los puntajes se acrecienta. El proceso de generación de puntajes de la teoría clásica⁷⁰ tiene un fundamento sencillo y comprensible para la población en general. La utilización de la teoría clásica impediría que dos estudiantes con el mismo número de respuestas correctas, omitidas e incorrectas tengan diferentes puntajes en la prueba. La situación anterior sería pan de cada día en el caso de asignarse puntajes utilizando IRT, y sería incomprensible para quienes no están familiarizados con la teoría.

Hay quienes sostienen que el hecho de que con IRT sea posible asignar distinto “peso” a preguntas con distinta dificultad y discriminación es una ventaja que por sí sola justificaría el uso de la teoría de respuesta al ítem en la asignación de puntajes. Sin embargo, de ser ése el único objetivo buscado, no se justifica el cambio de teoría. Con la teoría clásica es posible asignar distintos pesos a diferentes preguntas según criterios previamente establecidos, y el proceso de asignación de pesos en este caso podría hacerse más transparente a la opinión pública que al utilizar IRT. Sin embargo, aún en este último caso los beneficios que traería el asignar pesos diferenciados a las preguntas son cuestionables.

La asignación de puntajes por teoría clásica facilita también la tarea de los profesores que van a proveer a los estudiantes de pruebas de ensayo. Cuando se utiliza IRT, es muy difícil para un profesor predecir el puntaje que tendrá un alumno en la prueba real, ya que lo más seguro es que no tendrá los parámetros de las preguntas que él mismo ha diseñado para el ensayo, ni las herramientas para estimar el puntaje a partir de éstos. Cuando se utiliza teoría clásica la tarea de puntuar correctamente pruebas de ensayo se facilita sobremedida. No debemos perder de vista la experiencia internacional⁷¹ que nos enseña que estos problemas prácticos han sido una de las razones principales por las cuales se ha optado, siempre que ha sido posible, por mantener la asignación de puntajes por teoría clásica, especialmente en pruebas con consecuencias para los examinados. IRT se puede utilizar

⁷⁰ Porcentaje de respuestas correctas o número de correctas menos una fracción de las incorrectas, en el caso de preguntas de puntaje dicotómico.

⁷¹ Por ejemplo, el caso sueco y el caso de Estados Unidos.

para otros procesos paralelos en los cuales presenta ventajas reales, como estudios de sesgos, creación de bancos de datos, selección de preguntas y, en algunos casos, “equating”⁷².

Como se ha destacado anteriormente, IRT es un buen aporte a la evaluación educacional, pero no es la panacea. A continuación ahondaremos en este punto explicando los peligros de su sobrevaloración.

IRT es una teoría que se funda en una serie de supuestos que se cumplen rara vez en la realidad. Es importante evaluar que los supuestos se cumplan de manera adecuada, es decir, que si existen desviaciones del óptimo, éstas no sean tan importantes como para invalidar las aplicaciones de IRT y afectar significativamente la propiedad de invarianza de los parámetros. Para ello, es fundamental que expertos en el tema realicen las pruebas adecuadas para testear si efectivamente se está cumpliendo la unidimensionalidad y si el modelo se está ajustando a los datos experimentales. Si no es así, las medidas que se tomen al respecto deben ser compatibles con los objetivos primordiales de la prueba. Por ejemplo, si se detecta que un determinado tipo de pregunta es fuente de multidimensionalidad, se debe evaluar el efecto desde el punto de vista del contenido y propósitos de la prueba, y cuáles serían los efectos de su eliminación. Es posible que en muchos casos sea preferible cambiar el modelo estadístico a utilizar antes que desechar definitivamente el tipo de pregunta.

IRT es una teoría emergente que aún es debatida en el área académica. Existen muchos puntos a afinar y no hay consenso en cuáles son las pruebas óptimas para evaluar el cumplimiento de los supuestos y el ajuste del modelo a los datos. Es muy importante que los expertos encargados de implementar el modelo estén al tanto de las nuevas investigaciones y vayan avanzando conjuntamente con los progresos que se vayan dando en la teoría.

Es fundamental también que se estudie el efecto que tienen en los parámetros las variaciones significativas de experiencias educacionales de los estudiantes, variaciones de motivación y cambios en la ordenación de las preguntas en la prueba. De detectarse que hay efectos significativos en las estimaciones⁷³, deben tomarse medidas atinentes. Por ejemplo, si se detecta que un cierto tipo de preguntas presenta un marcado efecto de fatiga, debe cuidarse que la ubicación de éstas en las pruebas piloto sea similar a la ubicación que tendría en la prueba operacional. Debido a que es

⁷² Sin embargo, debe evaluarse para cada caso en particular cuál es el método idóneo de equating. Puede ser algún método basado en IRT, o alguno de teoría clásica.

⁷³ Es decir, que la invarianza no rige para poblaciones que han tenido diferentes experiencias educacionales, o que las preguntas presentan distintos parámetros según la posición que tienen en la prueba.

común que los parámetros estimados a partir de estudios piloto varíen, es altamente recomendable recalcularlos nuevamente en el contexto de la prueba operacional. Los parámetros obtenidos en la prueba piloto serán útiles para el proceso de selección de preguntas para la prueba, situación en la que es deseable, pero no fundamental, que éstos estén bien estimados. Sin embargo, para el cálculo final de los puntajes es primordial que los parámetros utilizados no estén sesgados⁷⁴, de manera de obtener una estimación acertada y precisa de la habilidad de los estudiantes examinados.

Es muy importante nunca perder de vista el objetivo primordial de la prueba. La sobrevaloración de los datos estadísticos que describen el comportamiento de las preguntas puede conducir a un actitud miope en el diseño de las pruebas, y traer consigo consecuencias desastrosas en la enseñanza. Es muy importante no sacrificar aspectos primordiales de la educación en aras del cumplimiento de los supuestos de una teoría. Por ejemplo, es muy posible que en una prueba de ciencias sociales se observen desviaciones significativas de la unidimensionalidad. Es probable que las preguntas de geografía estén midiendo un rasgo latente (o habilidad) diferente del que miden las preguntas de historia universal, o de economía. La solución estadística a este problema es simplemente seleccionar las preguntas de manera que se mantenga la unidimensionalidad. Eliminar las preguntas de geografía podría ser una solución en este caso particular. Lo anterior es razonable desde el punto de vista de IRT, pero no es razonable desde el punto de vista de los objetivos primordiales de la prueba. No debe perderse de vista que en una situación de medición el modelo debe adecuarse a los objetivos de la prueba y no viceversa. Una posible solución al problema de la multidimensionalidad y que no entraña tener que desechar preguntas, es la entrega de subpuntajes para cada uno de los rasgos independientes de la prueba. En el caso anterior esto se traduciría en entregar el puntaje de geografía separado del de historia. El único problema que esto podría traer es que estos subdominios pueden contener pocas preguntas, con lo que aumentaría el error de estimación de cada rasgo latente.

El problema de las violaciones de la unidimensionalidad no se da solamente en casos como el anterior, en que los subdominios corresponden claramente a contenidos diferentes. Puede suceder también que los distintos niveles de dificultad de las preguntas sean una fuente de multidimensionalidad. Por ejemplo, las preguntas más fáciles pueden estar midiendo el rasgo “capacidad de memorizar datos”, mientras las difíciles pueden estar midiendo “capacidad de razonamiento”. Ambas habilidades posiblemente serán

⁷⁴ Es decir, que no se desvíen de su valor real.

independientes y, por lo tanto, fuente de multidimensionalidad. También existe evidencia de que las preguntas de bloque, es decir, una agrupación de preguntas que se basan en un estímulo común (pasaje de lectura, gráfico, etc.), presentan dependencia local y por lo tanto son fuente de multidimensionalidad. Existe evidencia documentada de la multidimensionalidad del conocimiento científico. Shavelson y Lau (2002) hablan de esto y citan varios autores que han verificado la existencia de múltiples dimensiones en pruebas de conocimiento en estas áreas. Nuevamente hay que tener cuidado de no pasar a llevar los objetivos primordiales de la prueba en aras de mantener el supuesto de unidimensionalidad (por ejemplo, eliminando por completo las preguntas de razonamiento, eliminando los bloques de preguntas, o eliminando preguntas que reflejen alguna de las dimensiones del conocer científico).

Otro tema de radical importancia tiene que ver con la validez de las preguntas y la idoneidad editorial de éstas. La validez de las preguntas es un concepto bastante amplio, que tiene que ver con el nivel en que las preguntas son relevantes y representan efectivamente el campo que pretenden representar. La validez también tiene que ver con el nivel con que una pregunta está midiendo el constructo que intenta medir (por ejemplo, que una pregunta que pretende medir razonamiento matemático no esté midiendo en vez comprensión de lectura), y que sirva para el objetivo para el que se va a utilizar (si el objetivo de la prueba es predecir rendimiento académico en la universidad, que la habilidad medida sirva efectivamente para eso —validez predictiva—). La idoneidad editorial tiene que ver con la claridad en la diagramación y redacción de la pregunta. Los aspectos recién mencionados son tanto o más importantes que las teorías de medición utilizadas, y son totalmente independientes de ellas, es decir, muchos de los defectos de las preguntas en lo que se refiere a su validez de contenido e idoneidad editorial no son detectables por los métodos estadísticos de análisis de ítems. La sobrevaloración de las teorías de medición, exacerbada en el caso de IRT por ser una teoría tan promisoría, lleva a que en muchos casos se olvide que no basta con que una pregunta tenga la dificultad y discriminación adecuadas, sino que se debe cuidar de estar midiendo un conocimiento relevante en el contexto de los objetivos de la prueba.

Algo similar a lo anterior sucede cuando se tiene un exceso de confianza en las bondades de IRT en lo que se refiere al proceso de la construcción de las pruebas⁷⁵. Por ejemplo, en una prueba de currículum de biología las preguntas se escogerían de acuerdo al aporte que entregan en la

⁷⁵ Es decir, el proceso de seleccionar las preguntas desde un banco de ítems.

configuración de una función de información objetivo de la prueba⁷⁶. Luego se extrae del banco de ítems el grupo de aquellas que generan el mejor ajuste a ésta. Las preguntas contenidas en este grupo no necesariamente conformarán una muestra equilibrada del currículum. Por ejemplo, es posible que todas las preguntas seleccionadas tengan como tema la célula y ninguna trate el organismo humano, o ecología, o genética. A pesar de generar una curva de información idónea, el grupo de preguntas está lejos de cumplir los objetivos que tenía la prueba. Es por ello que los expertos en construcción de pruebas tienen que establecer en lo posible las restricciones adecuadas al proceso de selección de preguntas, y deben revisar acuciosamente el producto final obtenido.

No se debe nunca olvidar que IRT no reemplaza el criterio y sentido común en lo que se refiere a diseño de pruebas, sino que tan solo facilita y agiliza el proceso. No debe perderse de vista que tras el sofisticado aparataje matemático y computacional que trae consigo la teoría, aun estamos frente a un proceso de alta subjetividad, donde el criterio de quien lo aplica juega un rol fundamental.

REFERENCIAS BIBLIOGRÁFICAS

- Baker, F. *The Basics of Item Response Theory*. Estados Unidos: ERIC Clearinghouse on Assessment and Evaluation, segunda edición, 2001.
- Childs, R. y S. Oppler. "Implications of Test Dimensionality for Unidimensional IRT Scoring: An Investigation of a High Stakes Testing Program". *Educational and Psychological Measurement*, Vol. 60, N° 6 (2000).
- Dorans, N. "Item Parameter Invariance: The Cornerstone of Item Response Theory". *Research in Personnel and Human Resources Management*, Vol. 3 (1990).
- Fan, X. "Item Response Theory and Classical Test Theory: An Empirical Comparison of Their Item/Person Statistics". *Educational & Psychological Measurement*, Vol. 58, N° 3 (1998).
- Green, D.; W. Yen y G. Burket. "Experiences in the Application of Item Response Theory in Test Construction". *Applied Measurement in Education*, Vol. 2, N° 4 (1989).
- Hambleton, R.; H. Swaminathan y J. Rogers. *Fundamentals of Item Response Theory*. Sage Publications, 1991.
- Hamilton, L. S.; E. M. Nussbaum, H. Kupermintz, J. I. M. Kerkhoven y R. E. Snow, "Enhancing the Validity and Usefulness of Large-scale Educational Assessments: II. NELS: 88 Science Achievement". *American Educational Research Journal*, N° 32 (1995).

⁷⁶ Esta función de información objetivo es diferente según el tipo de prueba que se está diseñando y los objetivos que persigue. Por ejemplo, las funciones de información para una prueba de competencias mínimas deberán ser muy diferentes a las de una prueba de admisión a las universidades (véase sección 3.7).

- Hamilton, L. S.; E. M. Nussbaum y R. E. Snow. "Interview Procedures for Validating Science Assessments". *Applied Measurement in Education*, N° 10 (1997).
- Kingston, N. y N. Dorans. "Item Location Effects and Their Implications for IRT Equating and Adaptive Testing". *Applied Psychological Measurement*, Vol. 8, N° 2 (1984).
- Kirisci, L.; T. Hsu y L. Yu. "Robustness of Item Parameter Estimation Programs to Assumptions of Unidimensionality and Normality". *Applied Psychological Measurement*, Vol. 25, N° 2 (2001).
- Kolen, M. y R. Brennan. *Test Equating*. Springer-Verlag, 1995.
- Rudners, L. *Item Banking*. ERIC/AE Digest, 1998.
- Leary, F. y N. Dorans. "Implications for Altering the Context in Which Test Items Appear: A Historical Perspective on an Immediate Concern". *Review of Educational Research*, Vol. 55, N° 3 (1985).
- Linn, R. "Has Item Response Theory Increased the Validity of Achievement Test Scores?" *Applied Measurement in Education*, Vol. 3, N° 2 (1990).
- Lord, F. *Applications of Item Response Theory to Practical Testing Problems*. LEA Publishers, 1980.
- Loyd, B. "Implications of Item Response Theory for the Measurement Practitioner". *Applied Measurement in Education*, Vol. 1, N° 2 (1988).
- Nussbaum, E. M.; L. Hamilton y R. E. Snow. "Enhancing the validity and usefulness of large scale educational assessments: IV. NELS: 88 science achievement to 12th grade". *American Educational Research Journal*, N° 34 (1997).
- Shavelson, R. y S. Lau. "Multidimensional Validity Revisited: A Multidimensional Approach to Achievement Validation". *CSE Technical Report 574*. National Center for Research on Evaluation. University of California, Los Angeles, 2002.
- Stage, C. A. *Comparison Between Item Analysis Based on Item Response Theory and Classical Test Theory. A Study of the SweSAT Subtest ERC*. Umeå University, Department of Educational Measurement, 1998a.
- Stage, C. A. *Comparison Between Item Analysis Based on Item Response Theory and Classical Test Theory. A Study of the SweSAT Subtest WORD*. Umeå University, Department of Educational Measurement, 1998b.
- Stage, C. A. *Comparison Between Item Analysis Based on Item Response Theory and Classical Test Theory. A Study of the SweSAT Subtest READ*. Umeå University, Department of Educational Measurement, 1999.
- Tate, R. "Test Dimensionality". En G. Tindal y T. Haladyna (eds.), *Large-Scale Assessment Programs for All Students*. LEA Publishers, 2002 □.

**LA NECESARIA COMPLEMENTARIEDAD ENTRE
TEORÍA CLÁSICA DE LA MEDICIÓN (TCM)
Y TEORÍA DE RESPUESTA AL ÍTEM (IRT):
ASPECTOS CONCEPTUALES Y APLICACIONES***

Jorge Manzi y Ernesto San Martín

Este artículo tiene por objetivo mostrar la complementariedad conceptual y práctica que existe entre la Teoría Clásica de Medición (TCM) y la Teoría de Respuesta al Ítem (IRT). En la primera parte se presentan las similitudes estructurales entre ambos modelos, esto es, que ambas teorías se construyen a partir de dos hipótesis fundamentales: (1) que los instrumentos miden un rasgo unidimensional no observable, y (2) el axioma de independencia local. Luego se presentan los modelos IRT como una extensión natural de la TCM, la cual permite modelar el puntaje verdadero de un individuo no sólo en función de su rasgo latente propio, sino también por medio de las características de un ítem, como son su dificultad y su grado de discriminación. En la segunda parte del artículo se destacan las ventajas que se obtienen al analizar datos educacionales con un modelo estructuralmente más rico (es decir, los modelos IRT) en

JORGE MANZI. Psicólogo, P. Universidad Católica de Chile; Ph.D. en Psicología, Universidad de California, Los Angeles. Profesor Adjunto de la Escuela de Psicología de la Pontificia Universidad Católica de Chile.

ERNESTO SAN MARTÍN. Ingeniero Civil Matemático, Universidad de Chile; Ph.D. en Estadística, Universidad Católica de Lovaina, Bélgica. Profesor Auxiliar del Departamento de Estadística de la Pontificia Universidad Católica de Chile.

* Los autores han realizado consultorías para el SIMCE relativas a la comparabilidad de los puntajes de las pruebas y agradecen el acceso otorgado por el SIMCE a la información que ha sido empleada para ilustrar los temas abordados en este artículo.

comparación con la TCM. Dicha comparación se ilustra usando resultados de una prueba SIMCE aplicada a estudiantes de segundo año de enseñanza media. Los autores del artículo concluyen señalando que el enfoque IRT debe formar parte de la agenda de trabajo de todo esfuerzo serio en el ámbito de la medición educacional en gran escala.

1. Introducción

Las teorías de la medición son el fundamento conceptual que hace posible asociar números a los objetos que nos interesa medir. Por ello, han acompañado desde sus orígenes a las disciplinas científicas que fundan su quehacer en la medición. Esto es particularmente relevante en disciplinas como la psicología y la educación, que se interesan por medir aspectos que no son directamente observables, y que, por tanto, deben inferir atributos subyacentes a partir de los comportamientos observables de los examinados.

Las teorías de la medición tienen su origen hacia fines del siglo XIX. La formalización de la hoy llamada Teoría Clásica de la Medición (TCM) se remonta a comienzos del siglo XX. El desarrollo del análisis factorial fue una contribución complementaria que potenció dicha teoría y facilitó su aplicación para resolver problemas de gran interés científico, como ha sido la medición de inteligencia y personalidad. Luego de su gran dinamismo en la primera mitad del siglo XX, la TCM se constituyó en un cuerpo bien establecido de principios y procedimientos que han servido de base para el desarrollo de la mayoría de los más importantes instrumentos de medición que hoy conocemos. El texto de Gulliksen publicado en 1950 representa la consolidación del conocimiento acumulado por la TCM.

El dinamismo en torno a la medición se volcó en la segunda mitad del siglo pasado al desarrollo de modelos complementarios a la TCM, intentando superar los límites que se conocía tenía ese cuerpo de conocimientos. Dos fueron los desarrollos más relevantes: la teoría de la generalizabilidad (Cronbach, Gleser y Rajaratnam, 1963) y la Teoría de Respuesta al Ítem (IRT), también llamada teoría de rasgo latente. De ellas, esta última es la que ha concentrado marcadamente la atención académica a nivel internacional. Los trabajos pioneros de Rasch a fines de los años 50, y el impulso de Birnbaum y Lord en los años 60, sirvieron de base para el desarrollo de una teoría que hoy da cuenta de la mayor parte de la producción científica en este campo, y que comienza a ser considerada como el marco de referen-

cia básico para el diseño de nuevos instrumentos de medición. Es importante constatar que en este caso ha existido un importante rezago temporal entre el desarrollo de las ideas y sus aplicaciones prácticas, el que se explica fundamentalmente por las complejidades computacionales que supone el enfoque IRT. Sin embargo, a medida que se han desarrollado algoritmos y programas computacionales para resolver los problemas de estimación de parámetros y de habilidad de este modelo, su difusión ha comenzado a ampliarse. De hecho, una rápida revisión de revistas científicas no especializadas en materias psicométricas revela que el modelo IRT está siendo usado en múltiples contextos, tales como la medición de psicopatología (Rouse, Finger y Butcher, 1999; Waller, Thompson, y Wenk, 2000), personalidad (Fraleigh, Waller y Brennan, 2000; Gray-Little, Williams y Hancock, 1997; Panter, Swygert, Dahlstroim, Grant y Tanaka, 1997), inteligencia (Ellis, 1989; Godber, Anderson y Bell, 2000; Maller, 2000), etcétera.

Las primeras aplicaciones de este enfoque, que siguen siendo las más relevantes, se han concentrado en el campo educacional. Grandes programas de medición educacional a nivel internacional, tales como el TIMSS (matemática y ciencias), la medición de competencias lectoras (SIALS) o la evaluación de conocimientos y actitudes cívicas, se han apoyado en el enfoque IRT para desarrollar sus instrumentos y para reportar sus resultados. La medición del progreso educacional es también un área común, como lo muestra el uso de IRT en EE.UU. (NAEP) o en Chile (SIMCE). Más aun, en años recientes, se ha optado por este mismo enfoque para medidas educacionales individuales, como son la prueba norteamericana para seleccionar alumnos a programas de postgrado (el Graduate Record Examination, GRE) o la prueba que mide competencias lingüísticas en inglés (TOEFL).

Por ello, es relevante clarificar el rol y aporte que este enfoque representa. Como se describe en este trabajo, el enfoque IRT surge desde la TCM y representa una extensión de ella. No se trata, como algún observador ingenuo pudiera haber supuesto, de teorías alternativas o competitivas para explicar un mismo fenómeno. Por el contrario, se trata de teorías complementarias, que en la práctica deben ser usadas conjuntamente para resolver problemas básicos que toda medición representa.

Como se describe en este trabajo, el enfoque IRT busca resolver algunas limitaciones importantes de la TCM. En ciertos casos, esto ha llevado a desarrollos que no tienen paralelo en la Teoría Clásica (como la medición adaptativa) y, en otros, ha significado la ampliación de los recursos métricos para evaluar la calidad de un instrumento.

En este trabajo se representan las aplicaciones más establecidas y consolidadas del enfoque IRT, especialmente las que dicen relación con el análisis de instrumentos que emplean ítemes puntuados en forma dicotómica (que es lo usual en el contexto educacional). Para este tipo de aplicaciones existe, por lo demás, una variada oferta de programas computacionales. No consideraremos, en cambio, desarrollos y extensiones más recientes de este enfoque, tales como los modelos IRT multidimensionales, los modelos para ítemes puntuados en varias categorías de respuesta y los modelos no paramétricos (Van der Linden y Hambleton, 1997).

2. ¿Qué significa medir?

En nuestro país existe una amplia tradición de medición educacional. Las pruebas nacionales SIMCE y las pruebas de ingreso a la universidad son los principales ejemplos de ello. Las bases de datos generadas por dichas mediciones son utilizadas tanto para análisis rutinarios de medición, como para nuevos desarrollos teóricos y aplicados en el campo de la medición. Tanta actividad profesional y académica nos debe motivar a volver, cada cierto tiempo, sobre *cuestiones fundamentales*. Éstas, junto a la experiencia aplicada acumulada, nos permitirá tener una mejor comprensión de lo que significa medir y de los desafíos que ello implica.

¿Qué es medir? Parece una pregunta trivial. Sin embargo, si tratamos de explicar qué hacemos cuando medimos el largo de un lápiz, veremos que no es asunto fácil. Tanto más, si nos reducimos al fascinante campo de la medición educacional. En lo que sigue queremos revisar algunos elementos básicos de la *teoría de la medición*. Cuando medimos algún atributo de una clase de objetos o eventos (por ejemplo, medir el largo de un lápiz), *asociamos números (u otra entidad matemática familiar) con los objetos de una manera tal que las propiedades del atributo son completamente representadas por propiedades numéricas*. Supongamos que tenemos dos trozos rígidos de varillas, llamados *a* y *b*, cuyos largos queremos medir. Si ponemos una varilla al lado de la otra, entonces hay sólo tres posibilidades que pueden ocurrir: el extremo de la varilla *a* puede estar más allá del extremo de la varilla *b*, o el de *b* más allá del de *a*, o ambos extremos coinciden. En el primer caso decimos que *a* es más largo que *b*, en el segundo que *b* es más largo que *a*, y en el tercero que *a* y *b* tienen largos equivalentes. Por brevedad, escribimos $a > b$, $b > a$ o $a = b$, respectivamente.

Un primer elemento subyacente a toda medición es, por tanto, una *comparación cualitativa*; toda situación de medición involucra dicha com-

paración. Ejemplos más complejos son el de preferencias de canastas, grados de creencia frente a juegos de azar, mediciones astronómicas, etc. Un segundo aspecto fundamental subyacente a toda medición es la *evaluación cuantitativa de la comparación cualitativa*. El problema consiste, por tanto, en asignar números $\phi(a)$, $\phi(b)$, etc. a objetos a , b , etc. tal que el orden establecido cualitativamente se mantenga. Las evaluaciones numéricas se llaman *escala de medición*.

Un procedimiento natural para asignar números es el siguiente: asignamos a la primera varilla seleccionada un número *cualquiera*. Si el largo de la segunda varilla seleccionada excede el de la primera, le asignamos un número mayor; en caso contrario, uno menor. Hacemos lo mismo con una tercera varilla, excepto si su largo está entre las dos anteriores, en cuyo caso le asignamos un número que esté entre los dos anteriores. Este procedimiento puede ser realizado de manera indefinida. La simplicidad del ejemplo muestra que la asignación numérica (escala) es *arbitraria*, excepto que debe reflejar el orden cualitativo establecido u observado entre los objetos en cuestión.

Medir significa, por tanto, *asignar números tales que $\phi(a) > \phi(b)$ si y sólo si $a > b$* . No se trata sólo de definir una escala cualquiera tal que si $a > b$, entonces $\phi(a) > \phi(b)$, sino también que la escala sea tal que cuando $\phi(c) > \phi(d)$ entonces se pueda afirmar que $c > d$. Así, los problemas planteados por toda medición son los siguientes: (1) ¿qué propiedades debe satisfacer la relación de orden cualitativa que se observa o establece entre los objetos o atributos de una clase? Por ejemplo, puede requerirse un ordenamiento *total* de todos los elementos de la clase, de la misma manera en que están ordenadas las letras en el alfabeto; o tal vez podría tenerse una situación donde es imposible que todos los elementos de la clase sean comparables, por lo que sólo se puede definir un orden *parcial*; (2) ¿cómo asegurar la existencia de una escala de medición que preserve el orden cualitativo? Los resultados que se desean establecer dicen relación tanto con la *existencia* de una escala de medición (la función ϕ) como sus propiedades, todo lo cual depende esencialmente de las propiedades que el orden cualitativo satisface. Por ejemplo, si A es un conjunto de atributos, cuya cantidad de elementos es igual a todos los números naturales, y $<$ es un orden simple¹, entonces *existe* una escala ϕ sobre A tal que para todo a y b elementos de A , $a < b$ si y sólo si $\phi(a) < \phi(b)$. Digamos de paso que cuando se habla de escalas ordinales, escalas afines o escalas de intervalos, lo que se está requiriendo

¹ Esto es, que para todo a y b elementos de A , se tiene que $a < b$ o $b < a$; y además la relación $<$ es transitiva.

son ciertas propiedades que la función ϕ debe satisfacer. La escala ϕ satisface más propiedades cuanto más rica sea la estructura del orden cualitativo.

Estas consideraciones parecen suficientes para entrever lo complejo y fascinante que resulta el aplicar la teoría de medición en el ámbito de la medición educacional. Por lo pronto, baste decir que, si se quiere “medir”, el objetivo de un instrumento administrado a un grupo de estudiantes es ayudar a definir u observar una relación de orden cualitativa para *comparar alumnos*, estableciendo un ordenamiento entre ellos en base a un atributo o rasgo medido por un instrumento. El segundo problema, de no menor complejidad, consiste en encontrar (si es que existe) una escala que represente numéricamente dichas ordenaciones.

Para terminar esta sección, es importante mencionar que los primeros resultados de Teoría de Medición fueron establecidos a finales del siglo antepasado y principios del pasado; véase Cantor (1895), Helmholtz (1895) y Hölder (1901). A pesar de esto, la teoría de medición lleva desarrollándose de manera muy formal desde hace más de 60 años; sus principales resultados aún siguen apareciendo en el *Journal of Mathematical Psychology*. Para una exposición bastante completa y formal de la teoría, véase Pfanzagl (1968), Ellis (1968), Krantz *et al.* (1971) y Michell (1990). Para aspectos filosóficos y metodológicos, véase, entre otros, Swistak (1990), Hand (1996), Michell (1997a, 1997b).

3. ¿Cuáles son las hipótesis subyacentes a la TCM?

Cuando aplicamos un modelo estadístico a un conjunto de datos, un objetivo fundamental que se persigue es extraer *la información relevante contenida en ellos*. En la década de los 20, los métodos estadísticos enfatizaban la especificación y estimación de la distribución de probabilidades que generan los datos observados. Fisher (1922) escribió uno de los artículos seminales a partir del cual se desarrolló esta perspectiva. Sin embargo, las aplicaciones en econometría y psicometría motivaron una reformulación de los puntos de vista de Fisher. A principios de la década de los 50, en un artículo publicado en una de las mejores revistas de estadística matemática, *The Annals of Mathematical Statistics*, Koopmans y Reiersøl (1950) propusieron una reformulación del problema de la modelización tal y como lo concebía Fisher: “en muchos campos de aplicación, el objetivo de todo investigador no es sólo la población en el sentido de la distribución de probabilidades de las variables observables, sino también una estructura proyectada subyacente a esta distribución, por la cual esta última se asume

es generada” (p. 165). Estas ideas no sólo encontraron en Koopmans y Reiersøl (1950) su formulación formal, sino que motivaron todo un campo de investigación en psicometría y econometría; para detalles, véase Goldberger (1971, 1972), Novick (1980), Cox (1990), Manski (1995), McCullagh (2002) y Mouchart y San Martín (2003), San Martín (2003).

La Teoría Clásica de los Tests, así como la Teoría de Respuesta al Ítem, se han desarrollado teniendo como telón de fondo estas ideas. Los esfuerzos han estado concentrados en justificar, en la medida de lo posible, los modelos estadísticos con *elementos sustantivos*. Así, no sólo interesa la bondad de ajuste de los modelos estadísticos, sino también su significación sustantiva. El caso extremo sería un modelo que ajusta con un 99% de bondad, pero que su contenido sustantivo es tan pobre que no sirve para explicar los datos, ni mucho menos para aprender de ellos. Basta pensar, por ejemplo, en una regresión lineal con una cantidad bastante grande de variables explicativas.

En el campo de la medición cabe entonces preguntarse cómo se puede evaluar la significación sustantiva de un modelo estadístico, información que debe ser complementada con las más rutinarias como “buen ajuste estadístico”. Ya los antiguos psicometras que desarrollaron la Teoría Clásica de la Medición sugirieron ciertos lineamientos, los cuales han sido ampliamente desarrollados desde entonces. En efecto, no se trata de presentar un modelo estadístico, en este caso la TCM, listando un conjunto de supuestos o hipótesis, sino mostrar tanto su significación como sus implicaciones lógicas. Es necesario decir que estas consideraciones están ausentes de los actuales manuales de TCM (como de IRT) y, posiblemente, del uso profesional que se hace de los mismos. Sin embargo, si consultamos los trabajos originales, uno verá con sorpresa cómo estas consideraciones eran de gran importancia para el desarrollo de la teoría misma. Una de las referencias más importantes es Lord y Novick (1968), texto que no sólo desarrolla la TCM, sino también, y de manera integrada, los modelos IRT y los modelos factoriales². Otros trabajos que es importante mencionar aquí son Guttman (1945, 1953), Novick (1966) y Novick y Lewis (1967), entre otros. En esta sección queremos revisar dichos elementos, utilizando una notación más reciente a fin hacer explícita las relaciones estructurales que hay entre la TCM y los modelos IRT.

² Dicho libro fue escrito en colaboración con A. Birnbaum, uno de los estadísticos más conocidos por sus desarrollos en fundamentos de la Estadística. Birnbaum fue invitado a escribir los capítulos del libro que presentaron formalmente los fundamentos de la teoría de respuesta al ítem.

3.1. La estructura básica de la TCM

La TCM es, en apariencia, un modelo bastante simple. La variable dependiente corresponde al *puntaje total observado* de una persona particular. Las variables independientes son el *puntaje verdadero* y el *error de medición*. Se asume que estas variables están relacionadas *aditivamente*. Las hipótesis básicas del modelo son las siguientes: (a) el valor esperado del error de medición es cero; (b) el error de medición para una persona determinada no está relacionado con otras variables como el puntaje verdadero y el error de medición asociado a otros ítemes (véase Crocker y Algina, 1986; Gulliksen, 1950; Thissen y Wainer, 2001). Sin embargo, los “supuestos” (a) y (b) pueden ser deducidos de hipótesis más fundamentales, muchas de las cuales subyacen también a la Teoría de Respuesta al Ítem (IRT). En efecto, la TCM distingue dos procesos de aleatorios de selección: (i) la selección de los individuos, los cuales denotaremos por s ; (ii) la selección de los ítemes respondidos por un individuo s dado, los cuales son denotados por i . La primera hipótesis fundamental subyacente a cualquier teoría de medición es que las observaciones miden características individuales que no son directamente observables. Más precisamente,

Hipótesis 1: La respuesta del individuo s al ítem i mide un rasgo latente unidimensional, el cual caracteriza sólo al individuo s . Dicha variable la denotamos por θ_s .

Para hacer explícita la relación entre la observación (medición) con el constructo latente θ_s denotamos por X_{is} la respuesta del individuo s al ítem i . Entonces *siempre* podemos escribir la siguiente relación aditiva:

$$X_{is} = E[X_{is} | \theta = \theta_s] + (X_{is} - E[X_{is} | \theta = \theta_s])$$

donde $E[X_{is} | \theta = \theta_s]$ denota la esperanza condicional de X_{is} dado el constructo o variable latente θ_s . Esta esperanza condicional asume implícitamente que el individuo representado por θ_s ha sido escogido con respecto al proceso de selección (ii) mencionado arriba. En TCM, dicha esperanza condicional se llama *puntaje verdadero*, mientras que la diferencia $X_{is} - E[X_{is} | \theta = \theta_s]$ se llama *error de medición*. Ambas cantidades son *funciones de la variable latente* θ_s y, en consecuencia, no son observables directamente. Si $\tau_i(\theta_s)$ denotando el puntaje verdadero, y $\epsilon_i(\theta_s)$ denota el error de medición, la relación aditiva anterior se escribe equivalentemente como $X_{is} = \tau_i(\theta_s) + \epsilon_i(\theta_s)$. En el caso dicotómico (i.e. cuando cada ítem

tiene una y sólo una alternativa correcta) el puntaje verdadero se reduce a la probabilidad condicional de responder correctamente un ítem i dada la variable latente θ_s , esto es, $\tau_i(\theta_s) = P[X_{is}=1 \mid \theta = \theta_s]$.

Una pregunta relevante es la siguiente: ¿qué relación funcional hay que suponer para expresar el puntaje verdadero $\tau_i(\theta_s)$? Es evidente que esta pregunta no tiene una respuesta única. La TCM supone que esta es una relación lineal en θ asumiendo por tanto que τ es una función lineal y que todos los ítems de una prueba miden un mismo rasgo latente; véase Novick (1966). Hipótesis adicionales pueden ser introducidas a fin de obtener otras relaciones funcionales que deban ser ajustadas a los datos. Sin embargo, en lo que resta de esta sección, seguiremos haciendo la discusión independiente de la forma funcional, pues los resultados no dependen de la misma.

Usando las propiedades básicas de la esperanza condicional y sus relaciones con la covarianza³ es inmediato deducir que (i) $E[\epsilon_i(\theta_s)] = E[E(\epsilon_i(\theta_s) \mid \theta = \theta_s)] = 0$, y (ii) $\text{cov}(\tau_i(\theta_s), \epsilon_i(\theta_s)) = 0$. Estas propiedades son habitualmente presentadas como “supuestos” del modelo, sin embargo dependen de la hipótesis 1 ya introducida. Para obtener la propiedad según la cual la correlación entre los errores de medición debidos a dos ítems distintos es nula se obtiene a partir de la siguiente hipótesis, que no es otra cosa que el *axioma de independencia local*:

Hipótesis 2: Para cada individuo s , las mediciones X_{1s} , ..., X_{is} , ..., X_{ms} a los m ítems de una prueba son condicionalmente independientes dada la variable latente θ_s .

El axioma de independencia local afirma que las cantidades latentes (en el caso de la TCM, los puntajes verdaderos) son los únicos factores importantes, y que una vez que éstos han sido determinados, el comportamiento es aleatorio. Una excelente discusión heurística de este axioma puede encontrarse en Lazarsfeld (1959). Otras referencias importantes son Anderson (1959), Novick (1966), Holland y Rosenbaum (1980), Bartholomew (1987) y Sobel (1997). Así, usando otra propiedad fundamental de la covarianza⁴, se deduce que la hipótesis 2 implica que $\text{cor}(\epsilon_i(\theta_s), \epsilon_j(\theta_s))=0$ para dos ítems distintos i y j .

³ En particular, las siguientes dos propiedades: (i) si $E(X)$ existe, $E(X)=E[E(X \mid Y)]$; (ii) si $E(X)$, $E(Y)$ y $E(XY)$ existen, entonces $\text{cov}(X,Y)=\text{cov}[X,E(Y \mid X)]$.

⁴ A saber, si X , Y y Z son variables aleatorias tales que sus esperanzas existen, entonces $\text{cov}(X,Y)=E[\text{cov}(X,Y \mid Z)] + \text{cov}[E(X \mid Z),E(Y \mid Z)]$.

3.2. Algunas propiedades de los ítemes con respecto a la muestra

Cuando se analizan los ítemes de una prueba, interesa estudiar su comportamiento con respecto a una determinada muestra o *población de referencia*. Para la discusión que sigue, consideremos el caso dicotómico, siendo la extensión a otros casos inmediata. La dificultad media de un ítem i con respecto a una población de referencia de tamaño n se define por

$$X_i = \frac{1}{n} \sum_{s=1}^n X_{is}$$

Tanto el puntaje verdadero como el error de medición de un ítem dado pueden ser expresados en términos de una población de referencia. Para ello, es necesario considerar el proceso de selección aleatoria (ii) mencionado arriba. Así, el puntaje verdadero y el error de medición de un ítem i con respecto a una población de referencia están respectivamente dados por

$$\tau_i \equiv E[\tau_i(\theta)], \quad \tau_i(\theta) = \frac{1}{n} \sum_s \tau_i(\theta_s) \quad y \quad \varepsilon_i(\theta) = \frac{1}{n} \sum_s \varepsilon_i(\theta_s)$$

Las propiedades discutidas en la sección anterior se heredan a nivel poblacional. Por lo tanto, las hipótesis 1 y 2 anteriormente introducidas implican que, con respecto a una población de referencia, la TCM satisface las siguientes propiedades:

- (i) $X_i = \tau_i(\theta) + \varepsilon_i(\theta)$, para todo $i=1, \dots, m$.
- (ii) $E[\varepsilon_i(\theta)] = 0$, para todo $i=1, \dots, m$.
- (iii) $\text{cor}(\tau_i(\theta), \varepsilon_i(\theta)) = 0$, $i = 1, \dots, m$.
- (iv) $\text{cor}(\varepsilon_i(\theta), \varepsilon_j(\theta)) = 0$ para todo i distinto de j .

En particular, se deduce que $E[\tau_i] = E[X_i]$, esto es, que el puntaje verdadero esperado del ítem i es igual a su puntaje observado esperado.

Un “buen ítem” será aquel que tiene una “alta” correlación entre su puntaje observado X_i y su puntaje verdadero τ_i . De las propiedades anteriores se deduce que

$$\text{cor}^2(X_i, \tau_i(\theta)) = \frac{\text{Var}(\tau_i(\theta))}{\text{Var}(X_i)}.$$

Esta relación representa, por tanto, la precisión con que diferencias en el puntaje verdadero entre personas son estimadas por la diferencia entre

el puntaje observado entre personas. Por ello, esta relación se llama *confiabilidad del ítem i*. Mencionemos, sin embargo, que dicha confiabilidad no es calculable directamente pues depende de una cantidad no observable, a saber, θ . Es posible expresar la confiabilidad de un ítem bajo hipótesis adicionales, las cuales dicen relación con el concepto de medidas paralelas. Para detalles, véase Lord y Novick (1968, capítulo 3).

En conclusión, la TCM depende de dos hipótesis fundamentales: primero, de que las observaciones miden, con error, un rasgo latente, el cual a su vez caracteriza a cada individuo; y, segundo, del axioma de independencia local.

4. Los supuestos básicos de los modelos IRT

En la sección anterior, hemos mencionado el problema de cómo especificar el puntaje verdadero. La Teoría de Respuesta al Ítem ofrece respuestas a este problema. En efecto, suponiendo las hipótesis 1 y 2 asumidas por la TCM, la teoría IRT introduce hipótesis adicionales que permite *derivar relaciones funcionales específicas*. En efecto, los llamados *modelos 1PL* o *modelo Rasch* asumen las siguientes hipótesis adicionales:

Hipótesis 3: la función de distribución $P[X_{is}=1 \mid \theta = \theta_s]$ como función de θ es continua y estrictamente.

Hipótesis 4: El puntaje total de cada individuo obtenido en una prueba es un estadístico suficiente para θ_s . Es decir, el puntaje total contiene toda la información relevante provista por las observaciones $X1s, \dots, Xis, \dots, Xms$

Usando las hipótesis 1, 2, 3 y 4 se deduce que existen constantes $\beta_1, \dots, \beta_i, \dots, \beta_m$ tales que

$$P[X_{is} = 1 \mid \theta = \theta_s] = \frac{\exp(a(\theta_s - \beta_i))}{1 + \exp(a(\theta_s - \beta_i))}$$

Cuando $a=1$, este modelo se conoce como *modelo Rasch*. Para detalles, véase Fischer (1995a, 1995b) y Junker (2001). Como es sabido, los β_i 's representan la dificultad de cada ítem i . Así, el modelo 1PL puede ser visto como una extensión de la TCM en el sentido de introducir hipótesis adicionales a fin de especificar funcionalmente el puntaje verdadero de la TCM.

El modelo 1PL, y por tanto sus hipótesis subyacentes, implican una representación en una escala común del rasgo latente θ_s y de los parámetros de dificultad β_i de los ítemes. Más aún, ambos parámetros están representados en la *escala del logito*. En efecto, si denotamos por $P(X_{is})$ la relación funcional que define el modelo 1PL, entonces (para $a=1$, siendo los otros casos de misma interpretación)

$$\ln \frac{P(X_{is})}{1 - P(X_{is})} = \theta_s - \beta_i.$$

El cociente de probabilidades corresponde a la *razón de chance* que un individuo s tiene de responder correctamente el ítem i y de responder incorrectamente el mismo ítem. Dicha razón depende de dos características diferentes: la primera debida al individuo s (presente, por medio de la hipótesis 1, en la TCM), la segunda debida a los ítemes (presente en el modelo IRT gracias a las hipótesis 3 y 4). Cuando $\theta_s > \beta_i$, el individuo s tiene mayor *chance* de responder correctamente el ítem que de responderlo incorrectamente; la constante de proporcionalidad está dada por $\exp(\theta_s - \beta_i)$. Similarmente, cuando $\theta_s < \beta_i$, el individuo s tiene menor *chance* de responder correctamente el ítem que de responderlo incorrectamente. Hay un punto de inflexión cuando $\theta_s = \beta_i$, lo cual significa que la probabilidad de responder correctamente el ítem es igual a la probabilidad de responderlo incorrectamente; de hecho, ambas son iguales a 0.5.

De la ecuación de razón de *chance* ya mencionada, se puede deducir de forma inmediata que la diferencia entre el logaritmo de la razón de *chance* del individuo s con respecto al ítem i y con respecto al ítem j es igual a la diferencia $\beta_i - \beta_j$. En otras palabras, la escala en que se ordenan los ítemes es equivalente a la escala del logito en que se ordenan las razones de *chances*.

Una observación importante es la siguiente: para una persona s dada (luego, condicionalmente a θ_s), su puntaje total $X1s + X2s + \dots + Xms$ es utilizado en la TCM como una estimación de su puntaje verdadero en una prueba de largo infinito; véase Lord y Novick (1968, sección 5.4). En los cálculos que se hacen habitualmente en TCM, se reporta el puntaje total observado (o la proporción correspondiente). Siendo que el objetivo de cualquier teoría de medición es ordenar dos o más personas, la TCM ordena (compara) a los individuos de acuerdo a su puntaje total observado. Puesto que el puntaje total de la persona s en la prueba es un estadístico suficiente para el rasgo latente θ_s , el modelo 1PL *produce la misma ordenación entre los individuos que la producida por la TCM*. Desde este punto de vista, no hay diferencia alguna entre ambos modelos (o conjunto de hipótesis).

Todas estas propiedades estructurales del modelo son debidas a las hipótesis 3 y 4 añadidas a las hipótesis 1 y 2 usadas por la TCM. Como hemos ya ilustrado en los párrafos anteriores, esto permite obtener un modelo más rico en términos de interpretación que lo que podemos hacer con TCM.

Birnbaum desarrolló el llamado modelo 2Pl en el capítulo 17 de Lord y Novick (1968). La relación funcional que se asume es la siguiente:

$$P[X_{is} = 1 | \theta = \theta_s] = \frac{\exp(a_i(\theta_s - \beta_i))}{1 + \exp(a_i(\theta_s - \beta_i))}$$

Los parámetros α_i 's son interpretados como los parámetros de discriminación de cada ítem i . Este modelo también puede ser deducido a partir de hipótesis adicionales a las hipótesis 1, 2 y 3 mencionadas arriba. Referimos al lector interesado a Fisher (1987, 1994a). Un detalle importante que merece ser mencionado es que el puntaje ponderado es también un estadístico suficiente para el rasgo latente θ_s . Esto significa que todos los individuos teniendo el mismo puntaje ponderado deben tener también una misma estimación para θ_s . En particular, individuos con diferentes patrones de respuesta pueden tener el mismo puntaje ponderado. El modelo 2PL introduce entonces la importancia o “peso” que se les dan a los diferentes ítems de una prueba. Para otros detalles de estimación y discusión de los submodelos, véase Hambleton, Swaminathan y Rogers (1991) y Embretson y Reise (2000).

$$X_i = \frac{1}{n} \sum_{s=1}^n a_i X_{is}$$

Al igual que en el modelo 1PL, se puede deducir que la razón de *chance* de un individuo s con respecto a un ítem i es igual a $a_i(\theta_s - \beta_i)$. Así, el parámetro de discriminación a_i indica cuán mayor o menor es la probabilidad del individuo para responder correcta e incorrectamente el ítem. Dicho sea de paso que esto permite tener una caracterización alternativa de lo que significa el parámetro de discriminación.

5. Ventajas de los modelos IRT en relación a la TCM

Como hemos visto en las secciones anteriores, el enfoque IRT representa una extensión natural de la TCM, compartiendo algunos de sus supuestos fundamentales. Pese a sus evidentes similitudes con respecto a sus supuestos básicos, es importante apreciar también sus diferencias, pues

tanto en el plano conceptual como en sus aplicaciones prácticas, estas teorías poseen diferencias que es necesario considerar antes de resolver cuál de ellas es más pertinente para resolver qué tipo de problemas de medición. Estas diferencias son producidas por las hipótesis adicionales que el enfoque IRT asume, hipótesis que ayudan a extraer mayor información no sólo de los individuos, ni sólo de los ítems, sino también de sus interacciones.

5.1. La relación entre puntajes observados y rasgos latentes

Ambas teorías se fundan en afirmar una relación monotónicamente creciente entre los puntajes observados en un instrumento de medición y el atributo latente θ al mismo instrumento. Esta relación es conceptualmente razonable si se asume, como lo hacen ambas teorías, que el instrumento de medición evalúa fundamentalmente un atributo latente, esto es, se asume la unidimensionalidad del rasgo latente (véase hipótesis 1)⁵. Como fue discutido en las secciones anteriores, ambas teorías difieren en la forma funcional específica que postulan para la relación monótona creciente: la TCM plantea un supuesto más simple, sosteniendo que esta relación es lineal, mientras que el enfoque IRT plantea como un postulado fundamental que dicha relación es no lineal (véase hipótesis 3). En este último caso, un determinado instrumento de medición tiene la posibilidad de discriminar en ciertos rangos bajo y alto de la escala de θ . Por ello existe un comportamiento asintótico inferior y superior en los puntajes observados, si se considera un rango suficientemente grande de θ .

¿Cuál de los planteamientos parece más adecuado para modelar datos obtenidos en mediciones como las educacionales? Si bien es conceptualmente indiscutible que el planteamiento del enfoque IRT es más completo y consistente con lo que cabe esperar de la relación entre θ y puntajes observados⁶, es evidente que se trata, al mismo tiempo, de un modelo más complejo. Por otro lado, el planteamiento de la TCM es evidentemente más simple y, por lo mismo, más parsimonioso. Esta competencia entre precisión y parsimonia explica en gran medida la preferencia que en términos prácticos llevan a emplear uno u otro modelo. Quienes privilegian la sim-

⁵ En ambas teorías la evaluación de este supuesto se lleva a cabo fundamentalmente con el apoyo del análisis factorial exploratorio. Para probar que un instrumento es unidimensional usualmente se compara la magnitud del valor propio asociado al primer factor con el de los restantes valores propios.

⁶ Es relevante mencionar aquí que las discusiones de Rasch en cuanto a la pertinencia de los modelos IRT se basaron no sólo en caracterizaciones matemáticas, sino también en experimentos sustantivos. Para un excelente resumen acerca de cómo surgió el modelo Rasch, véase Andersen y Olsen (2001).

plicidad, probablemente preferirán el supuesto lineal. Quienes, en cambio, prefieren un modelo que resuelva en forma conceptualmente más precisa el problema planteado al comienzo de esta sección, preferirán IRT.

¿Qué nos muestra la realidad? Es útil ilustrar el dilema planteado a partir de datos empíricos basados en aplicaciones prácticas de las teorías de la medición. Con este propósito, trabajaremos con datos obtenidos de la aplicación de la prueba SIMCE, puesto que, por su carácter censal, se dispone de información que potencialmente refleja la realidad de toda la población de interés. En términos concretos, la mayoría de los ejemplos que se discutirán en el presente trabajo se basan en datos provenientes de la última aplicación de la prueba SIMCE de matemáticas a estudiantes de 2° medio, utilizando una de las formas de dicha prueba. Para el problema que nos interesa, la relación entre atributo latente y puntaje observado, mostraremos en primer lugar la aproximación más básica a dicha relación que puede derivarse de la TCM: la relación entre puntaje total en la prueba y la probabilidad de responder correctamente cada uno de los ítemes que conforman dicha prueba. Como se indicó más arriba, la TCM emplea como estimador de θ el puntaje total en un conjunto de ítemes (la suma o promedio de las respuestas que los examinados producen en el conjunto de ítemes). Para ilustrar la relación, se agruparon todos los examinados en rangos de 4 puntos en el total de ítemes⁷; luego se calculó el promedio de respuesta correcta de cada grupo para cada uno de los 45 ítemes de la prueba. Tal como puede apreciarse en los gráficos de la Figura 1 del Anexo, todos los ítemes muestran un comportamiento monotónicamente creciente con respecto al puntaje total. Adicionalmente, se observa que la mayoría de ellos presenta una relación no lineal con el puntaje total: en algunos casos (como el ítem 6), con una evidente asíntota inferior, en otros (como el ítem 1), con una clara asíntota superior, y en otros casos con una forma cercana a una ojiva. Todos estos patrones son consistentes con los planteamientos del enfoque IRT.

5.2. Los supuestos acerca del error de medición

Como discutimos en la sección 3.1., toda teoría de la medición asume que las respuestas observadas miden, con error, los atributos. Los errores de medición se definen con respecto a una población (como lo desarrollado en la sección 3.2.) o con respecto a un instrumento o test. En

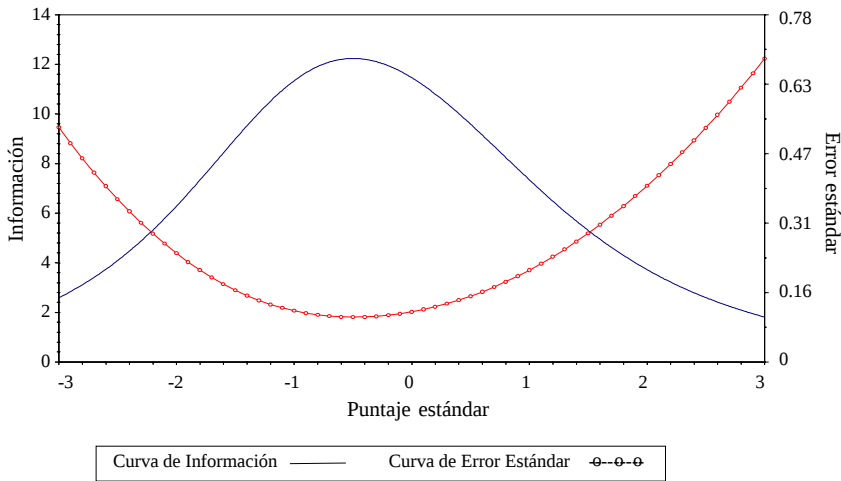
⁷ Para estimar los puntajes totales de dicha prueba se emplearon 45 ítemes. Por tanto, los puntajes observados podían variar entre 0 y 45 puntos. Para agrupar los examinados, se crearon intervalos de 4 puntos en el puntaje total. La primera categoría, sin embargo, abarca de 0 a 9 puntos, para asegurar que un número mínimo de examinados quedara en cada categoría. Ninguna categoría incluye menos de 1.200 casos.

este último caso, los supuestos de la TCM implican que dicho error es constante para un mismo instrumento, transformándose así en un atributo que caracteriza la calidad de la medición y de las inferencias que pueden hacerse a partir del instrumento; véase Lord y Novick (1968, capítulo 3). Sin embargo, la intuición sugiere que dicho error difícilmente es constante: basta conocer cómo se construyen los instrumentos de medición para concluir que la precisión de la medición seguramente es desigual. Por ejemplo, si se está construyendo una prueba de carácter selectivo, seguramente se seleccionarán más preguntas difíciles que fáciles, con lo que la precisión debiera ser mayor en la región superior de los puntajes que en la inferior. Lamentablemente, la TCM emplea un supuesto simplificador, que en este caso puede resultar peligroso, puesto que lleva a muchos usuarios de instrumentos de medición a pensar que la calidad con que se está evaluando es constante, ignorando que el error de medición puede ser sensiblemente más alto en ciertas regiones de los atributos evaluados.

Los supuestos subyacentes al enfoque IRT permiten subsanar esta situación, incorporándose la noción de error no constante; en efecto, se deduce que dicho error varía de manera inversamente proporcional a la información que un instrumento de medición posee. Así, tal como puede observarse en la Figura 2, el error, siguiendo lo que nuestra intuición sugiere, es mayor en las regiones extremas de la escala de medición, y menor en la zona media. Esto refleja además el hecho que hay menos individuos con patrones de respuestas en las zonas extremas que en las zonas medias. Es importante tener presente que no necesariamente un instrumento debe mostrar el patrón que aparece en la Figura 2. De hecho, se puede construir un instrumento que resulte especialmente informativo en la región superior o inferior de un determinado atributo, si las aplicaciones prácticas del instrumento así lo aconsejen. Cómo resolver el problema de diseñar un instrumento de medición para satisfacer propósitos específicos, es una tarea para la cual IRT provee herramientas conceptuales y técnicas precisas. De hecho, aplicaciones de creciente importancia en la medición internacional, tales como la medición adaptativa o la medición basada en computadores (computer based testing), se fundan en las reglas del escalamiento óptimo que ofrece IRT⁸. Para este tipo de aplicaciones, la TCM no ofrece reglas ni principios que puedan resolver en forma satisfactoria y precisa el problema. De hecho, la única regla claramente deducible de la TCM cuando se desea incrementar la precisión de un instrumento, es incrementar su longitud, basándose en la profecía de Sperman-Brown (véase Lord y Novick, 1968;

⁸ El escalamiento óptimo alude a la selección de ítemes que otorgan el máximo de información y, en consecuencia, el mínimo de error, en determinada región del atributo que interesa evaluar.

FIGURA 2: FUNCIÓN INFORMATIVA Y ERROR ESTÁNDAR DE MEDICIÓN DE LA PRUEBA SIMCE DE MATEMÁTICA



Crocker y Algina, 1986), lo que en muchos casos lleva a soluciones impracticables. En cambio, como lo hace ver Embretson y Reise (2000), IRT provee indicaciones directas para aumentar la precisión de un instrumento, sin que ello conlleve una extensión de su longitud.

5.3. Dependencia entre estimaciones de habilidad e ítemes de una prueba

Las teorías de la medición se basan en el supuesto que una mejor aproximación al puntaje verdadero de una persona se obtiene cuando se emplean varios indicadores o ítemes. En estas condiciones, lo deseable sería que, para un mismo número de ítemes, las estimaciones de habilidad que se obtengan sean relativamente independientes de los ítemes específicos que se empleen. Lamentablemente esta segunda condición no está incorporada en la TCM. En este enfoque, la estimación de habilidad está intrínsecamente ligada a los ítemes específicos que conforman una prueba. Si se alteran o cambian los ítemes, aunque los nuevos midan el mismo rasgo o atributo latente, se obtienen distintas estimaciones de habilidad. En términos prácticos, es fácil deducir que si de un conjunto de ítemes se extrae una muestra de un cierto tamaño de ítemes fáciles y otra de igual tamaño de ítemes difíciles, la mayor parte de los examinados obtendrá un

mayor puntaje (y por tanto, estimación más alta de su habilidad) si se le administra el primer conjunto de ítemes. En cambio, el enfoque IRT incorpora en la estimación de habilidad la calidad de los ítemes que han sido abordados y respondidos. Así, una persona que sólo responde correctamente un subconjunto de preguntas fáciles obtendrá en este enfoque un menor puntaje que una persona que responde correctamente un mismo número de preguntas, si estas últimas son más difíciles. Desde este punto de vista, la selección específica de los ítemes que conforman una prueba es una decisión que tiene consecuencias más críticas en la TCM que en el enfoque IRT.

Lo anterior tiene, nuevamente, importantes consecuencias en el diseño de pruebas. Para obtener una estimación adecuada de la habilidad de una persona, en el enfoque de la TCM se requiere que quienes son examinados sean expuestos a una muestra más o menos representativa del conjunto de ítemes que es posible concebir para el dominio que se está evaluando. En términos prácticos, esta condición es muy difícil de satisfacer, puesto que las pruebas sólo pueden contener un número relativamente limitado de ítemes o reactivos. Por lo demás, la TCM tampoco provee reglas precisas para establecer si se dispone de un banco de preguntas relativamente representativo. En cambio, empleando las reglas del enfoque IRT, es menos crítico cuáles ítemes específicos sean administrados a un examinado. Si los parámetros de los ítemes que se administran han sido calibrados en forma conjunta, una muestra relativamente pequeña de preguntas permite obtener estimaciones adecuadas de habilidad, y por lo demás comparables entre distintos examinados, aun si distintas personas son expuestas a diferentes preguntas.

5.4. Dependencia de los parámetros de las muestras en que son estimadas

Según lo visto hasta aquí, un aspecto fundamental de la construcción de instrumentos de medición es la calidad de los ítemes que conforman una prueba. Tanto la TCM como el enfoque IRT contemplan parámetros para evaluar la calidad de los ítemes disponibles. Los parámetros más importantes en la tradición de la TCM son los que aluden al grado de dificultad⁹ de

⁹ El término grado de dificultad es algo equívoco, puesto que es sólo pertinente para ítemes que miden habilidades cognitivas. Cuando se miden otros aspectos, como por ejemplo, actitudes, el promedio de respuesta de un ítem no puede interpretarse como grado de dificultad. Por otra parte, cuando la escala de respuesta no es dicotómica, el promedio de un ítem tampoco corresponde a una proporción de respuesta correcta, como es usualmente identificado el grado de dificultad. En consecuencia, el concepto más general sería la posición promedio del ítem en la escala de respuesta que se esté empleando.

un ítem y a su capacidad discriminativa. En este enfoque, un ítem es apropiado en la medida que su grado de dificultad se encuentra dentro de un rango aceptable (típicamente entre 0,2 y 0,8, o entre 0,1 y 0,9), y en la medida que su capacidad discriminativa sea adecuada (habitualmente esto se especifica con exigencias mínimas para la correlación entre el puntaje en el ítem y el puntaje en la prueba total¹⁰). En el enfoque IRT, las hipótesis adicionales subyacentes al modelo muestran que las características de un ítem son fundamentales para definir el puntaje verdadero de un individuo (o probabilidad de responder correctamente un ítem). En efecto, como vimos en la sección 4, el modelo de 1 parámetro (o modelo Rasch) incluye la posición de cada ítem (que tiene una interpretación análoga al grado de dificultad en la TCM). Más aún, el modelo de 1 parámetro también considera la capacidad discriminativa de los ítems, la cual se asume constante para todos ellos. El modelo de 2 parámetros, por su parte, considera que tanto la posición como la capacidad discriminativa de los ítems son atributos específicos para *cada* ítem. Mencionemos que existe un tercer modelo que involucra 3 parámetros; la relación funcional entre la variable observada y el atributo latente es tal que se acepta la posibilidad que la asíntota inferior de dicha función no se sitúe en 0. Típicamente esto último ocurre en ítems de respuesta cerrada, donde existe la probabilidad de responder correctamente al azar.

Lo deseable sería que la estimación que se obtuviera de los parámetros de los ítems pudiera ser relativamente independiente de las características de las muestras involucradas en su estimación. Lamentablemente ello no puede ser garantizado en la TCM, puesto que por definición el grado de dificultad de un ítem está definido como el promedio que obtienen en él quienes responden dicho ítem. Por tanto, si la muestra incluye a personas con un mayor nivel de habilidad, la proporción de respuestas correctas debiera ser, por definición, mayor que si el ítem es respondido por un grupo de examinados con menor nivel de habilidad. En consecuencia, la dificultad no es en este caso un atributo del ítem, sino que una interacción entre el ítem y los examinados que lo responden. Con respecto a la capacidad discriminativa, se puede asumir que será relativamente menos dependiente que el grado de dificultad de las características de la muestra, condicionado al grado de variabilidad que exista en la muestra. En contraste con lo anterior, el enfoque IRT separa la información debida a un individuo y la información debida a un ítem; esto permite decir (como fue discutido ya en la sección 4) que los parámetros que caracterizan los ítems son relativa-

¹⁰ Cuando los ítems son puntuados dicotómicamente esta correlación adopta la forma de una correlación biserial o de una correlación biserial-puntual.

TABLA 1: PARÁMETROS TCM E IRT DE 45 ÍTEMES DE PRUEBA SIMCE DE MATEMÁTICAS

Ítem	PropCorrecta	Desv Est	Corr _{it-test}	a _{j(todos)}	ErrEst(a _j)	b _{j(todos)}	ErrEst(b _j)
m1	0.80	0.40	0.28	0.64	0.01	-1.55	0.02
m2	0.48	0.50	0.38	0.60	0.01	0.06	0.01
m3	0.59	0.49	0.32	0.52	0.01	-0.54	0.01
m4	0.58	0.49	0.37	0.63	0.01	-0.42	0.01
m6	0.21	0.40	0.31	0.49	0.01	1.83	0.02
m7	0.82	0.39	0.22	0.46	0.01	-2.12	0.04
m8	0.74	0.44	0.38	0.89	0.01	-0.97	0.01
m9	0.55	0.50	0.41	0.70	0.01	-0.29	0.01
m10	0.43	0.50	0.29	0.40	0.01	0.41	0.01
m11	0.57	0.49	0.37	0.62	0.01	-0.39	0.01
m12	0.48	0.50	0.30	0.45	0.01	0.10	0.01
m13	0.83	0.37	0.20	0.45	0.01	-2.31	0.04
m14	0.64	0.48	0.38	0.70	0.01	-0.66	0.01
m15	0.51	0.50	0.42	0.72	0.01	-0.11	0.01
m16	0.47	0.50	0.34	0.50	0.01	0.14	0.01
m17	0.49	0.50	0.33	0.50	0.01	-0.01	0.01
m18	0.42	0.49	0.38	0.59	0.01	0.31	0.01
m19	0.61	0.49	0.37	0.64	0.01	-0.56	0.01
m20	0.85	0.36	0.27	0.73	0.01	-1.70	0.02
m21	0.75	0.43	0.40	1.05	0.01	-0.95	0.01
m22	0.72	0.45	0.40	0.98	0.01	-0.87	0.01
m23	0.46	0.50	0.46	0.80	0.01	0.09	0.01
m24	0.48	0.50	0.37	0.56	0.01	0.06	0.01
m26	0.30	0.46	0.25	0.37	0.01	1.46	0.02
m27	0.36	0.48	0.29	0.42	0.01	0.83	0.02
m28	0.26	0.44	0.24	0.34	0.01	1.94	0.03
m29	0.58	0.49	0.42	0.78	0.01	-0.41	0.01
m30	0.63	0.48	0.44	0.91	0.01	-0.56	0.01
m31	0.51	0.50	0.37	0.59	0.01	-0.08	0.01
m32	0.64	0.48	0.39	0.76	0.01	-0.64	0.01
m33	0.50	0.50	0.39	0.60	0.01	-0.07	0.01
m34	0.61	0.49	0.43	0.83	0.01	-0.48	0.01
m35	0.34	0.47	0.40	0.63	0.01	0.72	0.01
m36	0.38	0.49	0.48	0.84	0.01	0.38	0.01
m37	0.37	0.48	0.34	0.52	0.01	0.67	0.01
m38	0.36	0.48	0.52	0.92	0.01	0.42	0.01
m39	0.38	0.49	0.36	0.56	0.01	0.56	0.01
m40	0.40	0.49	0.24	0.35	0.01	0.73	0.02
m42	0.53	0.50	0.38	0.41	0.01	2.22	0.03
m43	0.67	0.47	0.27	0.62	0.01	-0.21	0.01
m44	0.80	0.40	0.34	0.44	0.01	-1.09	0.02
m45	0.47	0.50	0.37	0.91	0.01	-1.21	0.01
m46	0.42	0.49	0.23	0.58	0.01	0.09	0.01
m47	0.31	0.46	0.16	0.32	0.01	0.60	0.02
m48	0.43	0.50	0.41	0.23	0.01	2.05	0.05
Promedio	0.53	0.47	0.35	0.61	0.01	-0.06	0.01

mente independientes del grupo de examinados que se emplee para estimarla.

Para ilustrar este aspecto, emplearemos nuevamente los datos del SIMCE como marco de referencia. La Tabla 1 contiene para cada uno de los 45 ítemes sus parámetros TCM e IRT estimados con todos los casos. Para evaluar la estabilidad de ellos en submuestras, se obtuvieron muestras al azar de tamaño algo superiores a los 2.000 estudiantes bajo 2 condiciones: muestras del conjunto de todos los examinados y muestras de grupos que poseen importantes diferencias de puntaje, como es el caso de quienes asisten a establecimientos de distinta dependencia¹¹ (municipalizados, particulares subvencionados y particulares pagados). En cada caso se estimaron los parámetros TCM e IRT y luego se compararon sus valores con los estimados en el conjunto de la población.

Los resultados, que se resumen en la Tabla 2, muestran con claridad que cuando se emplean muestras al azar, se obtienen estimaciones muy cercanas a los valores poblacionales tanto de los parámetros IRT como de

TABLA 2: CORRELACIONES ENTRE PARÁMETROS TCM E IRT DE LOS ÍTEMES ESTIMADOS EN DISTINTOS SUBGRUPOS
(Las correlaciones incluyen 45 ítemes de una de las formas de la prueba SIMCE de Matemática.)

	Grado de dificultad		Capacidad discriminativa	
	TCM (Pj)	IRT (bj)	TCM (rit-test)	IRT (aj)
Población - Aleatoria 1	1.00	1.00	.98	.97
Población - Aleatoria 2	1.00	1.00	.98	.98
Población - Particulares	.86	.96	.67	.63
Población - Subvencionados	1.00	.99	.95	.95
Población - Municipalizados	.99	.99	.92	.94
Particulares - Subvencionados	.87	.94	.54	.54
Particulares - Municipalizados	.82	.93	.45	.43
Subvencionados - Municipalizados	.98	.99	.95	.95

Población: estimaciones basadas en todos los estudiantes del país que rindieron la prueba.

Aleatoria 1 y Aleatoria 2 son dos muestras independientes obtenidas al azar.

Particulares, Subvencionados y Municipalizados corresponden a muestras obtenidas al azar de estudiantes de cada una de las dependencias.

Todas las muestras incluyen unos 2.200 estudiantes.

¹¹ Es importante considerar que estos 3 tipos de establecimientos poseen tamaños desiguales en el país, lo que explica que los valores poblacionales se asemejen más a los grupos de mayor tamaño (los municipalizados representan el 47,5% del total de estudiantes, los subvencionados el 42,9% y los particulares el 9,6%).

los parámetros de la TCM. Sin embargo, cuando se emplean grupos con diversa habilidad como base para estimar los parámetros, se obtienen resultados más divergentes. En concordancia con los supuestos teóricos, el grado de dificultad es menos consistente al estimarlo con TCM que al hacerlo con IRT. Por su parte, las estimaciones de la capacidad discriminativa de los ítems muestran correlaciones relativamente semejantes, aunque algo inferiores a las obtenidas para los parámetros asociados al grado de dificultad.

5.5. Representación simultánea de habilidades y dificultades

Como hemos mencionado en la sección 4, los modelos IRT permiten una representación simultánea de habilidades y dificultades. Esto representa sin duda una ventaja adicional del modelo IRT sobre el de la TCM. En efecto, en la teoría clásica, las habilidades de los individuos y las dificultades de los ítems son representadas en escalas diferentes: el grado de dificultad de un ítem se expresa (habitualmente) como una proporción de respuesta correcta, mientras que la habilidad de las personas se expresa como un puntaje total (en su forma bruta, como suma o promedio de respuestas correctas, o en su forma estandarizada, luego de transformar los puntajes a otra escala).

Esta representación simultánea habilidad-dificultad producida por los modelos IRT ofrece ventajas interpretativas adicionales. En efecto, gracias a ella es posible referir el rendimiento de un examinado al tipo de ítems cuya localización está por debajo o por encima de él; esto, a su vez, facilita la posibilidad de construir estándares para interpretar los resultados de una prueba. Por otra parte, cuando la ubicación de los ítems está en la misma escala que los examinados, es posible identificar con facilidad el tipo de ítems que resultaría más informativo para evaluar a una determinada persona. Este es el principio que orienta la medición adaptativa, especialmente en su implementación secuencial, donde usualmente con el apoyo de un computador, se determina en cada etapa de una medición cuál sería el ítem que resultaría más informativo para estimar la habilidad de una persona, a partir del rendimiento que dicha persona haya obtenido en los ítems que se le hayan administrado previamente. Esta propiedad, que no tiene paralelo en la TCM explica por qué el enfoque IRT es la única teoría de la medición que puede ser empleada como fundamento para la medición adaptativa. Este tipo de aplicaciones ya forma parte de mediciones educacionales en gran escala, como son la prueba para medir competencias lingüísticas

en inglés (TOEFL), o la prueba de selección para estudios de postgrado en EE.UU. (Graduate Record Examination).

5.6. Comparación de puntajes entre distintas pruebas (*equating*)

Finalmente, cabe mencionar que el enfoque IRT facilita otra importante aplicación de la teoría de la medición: la posibilidad de hacer comparables las puntuaciones de dos o más instrumentos de medición. Esta capacidad fue implícitamente aludida previamente al mencionar que en este enfoque las puntuaciones que obtienen las personas no dependen de los ítems específicos que se le administran a una determinada persona. En concreto, si un conjunto de ítems ha sido calibrado en forma conjunta, o sus parámetros han sido establecidos en una misma escala, se pueden construir diversas formas o pruebas a partir de dichos instrumentos, cuyos puntajes estarían expresados en una misma escala. Desde este punto de vista, la comparabilidad de puntuaciones (*equating*) es consustancial al modelo IRT. En cambio, en la TCM la posibilidad de realizar comparaciones entre puntajes de personas que han respondido diversas formas de una misma prueba, requiere el empleo de diseños y procedimientos especiales, puesto que, como se precisó antes, las puntuaciones en este enfoque no son comparables, salvo en el caso que se basen en formas estrictamente paralelas de un mismo instrumento.

6. Aplicaciones prácticas del enfoque IRT

Hasta aquí se ha aludido a las semejanzas y diferencias básicas entre los modelos TCM e IRT. En esta sección se ilustrará cómo puede aplicarse en forma práctica el enfoque IRT para resolver los problemas habituales de medición. Las aplicaciones prácticas de la teoría clásica son más conocidas y están por lo demás disponibles en software estadístico general, como SPSS, SAS o STATISTICA. IRT, en cambio, representa un modelo matemáticamente más complejo, cuya aplicación supone el empleo de técnicas y programas especializados. A continuación se ilustrará, recurriendo al ejemplo de la prueba SIMCE de matemática, algunas aplicaciones prácticas del modelo IRT.

6.1. Determinación de la dimensionalidad de los ítems

El supuesto de unidimensionalidad, como fue mencionado previamente, constituye la base de la mayor parte de los modelos IRT. Aunque en

años recientes se han desarrollado modelos multidimensionales, e incluso se está desarrollando software para hacer posible su implementación, las aplicaciones prácticas de este enfoque suponen unidimensionalidad. Para evaluar este supuesto se puede recurrir a diversas técnicas (véase Hambleton, Swaminathan y Rogers, 1991; Embretson y Reise, 2000). La más relevante de ellas es el análisis factorial exploratorio. Con esta técnica se busca determinar el número de dimensiones que subyacen a un conjunto de ítems, a partir del análisis de la matriz de intercorrelaciones entre ellos. Un problema práctico que aparece al aplicar esta técnica a ítems puntuados dicotómicamente, es que la correlación regular de Pearson entre variables dicotómicas (usualmente denominada coeficiente Phi) no se acomoda al modelo factorial. Este tipo de correlaciones resulta atenuada con respecto a las que se obtendrían si las variables fueran continuas, con lo que se subestiman los pesos factoriales. Por otra parte, cuando las probabilidades de responder correctamente los ítems difieren entre ellos, que es lo usual, el modelo factorial convencional sobreestima el número de factores. Finalmente, los ítems dicotómicos no se relacionan en forma lineal con las dimensiones subyacentes continuas (factores). Por todas estas razones, la recomendación tradicional ha sido efectuar análisis factorial de ítems a partir de correlaciones tetracóricas entre los ítems (Mislevy, 1986; Woods, 2002). Por desgracia esta opción no está implementada en programas estadísticos convencionales, por lo que es necesario recurrir a programas ad hoc, como TESTFACT (Bock *et al.*, 2003). Una opción a dicho programa es Mplus (Muthén y Muthén, 1998), un programa especializado en sistemas de ecuaciones estructurales (SEM), que ha incorporado opciones para variables categóricas, lo que permite manejar ítems dicotómicos.

Para el caso de la prueba SIMCE de matemática, se llevó a cabo el análisis factorial implementado en TESTFACT 4.0, calculando la matriz de correlaciones tetracóricas entre los ítems. Los resultados demostraron, en primer lugar, que había razonable evidencia de unidimensionalidad en esta prueba, tal como lo atestigua el gráfico de los eigenvalues (Figura 3). En él se puede apreciar que el primer factor se destaca muy claramente de los restantes. Adicionalmente, los pesos factoriales de los ítems en el primer factor demuestran que el primer factor se correlaciona en forma sustantiva con casi todos los ítems. Sólo los ítems 5 y 25 muestran bajas correlaciones, por lo que se decidió excluirlos de los análisis subsecuentes¹². En suma, esta parte de los análisis indica que la estructura de esta prueba es consistente con los requerimientos del enfoque IRT.

¹² Estos dos ítems, así como el 41 (por razones que se aclaran en la siguiente sección) fueron finalmente excluidos de la calibración de la prueba de matemática. Ello determinó que en definitiva se emplearan 45 de los 48 ítems de esta forma de dicha prueba. Los análisis reportados en secciones anteriores de este trabajo habían excluido estos ítems.

FIGURA 3: GRÁFICO DE EIGENVALUES BASADO EN EL ANÁLISIS FACTORIAL DE LA PRUEBA SIMCE DE MATEMÁTICA

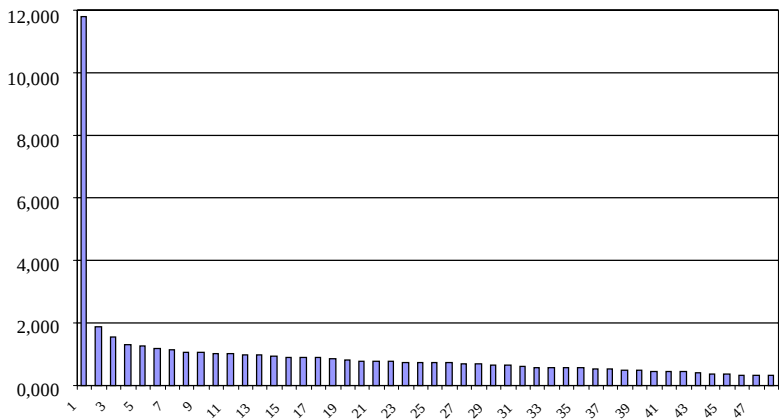


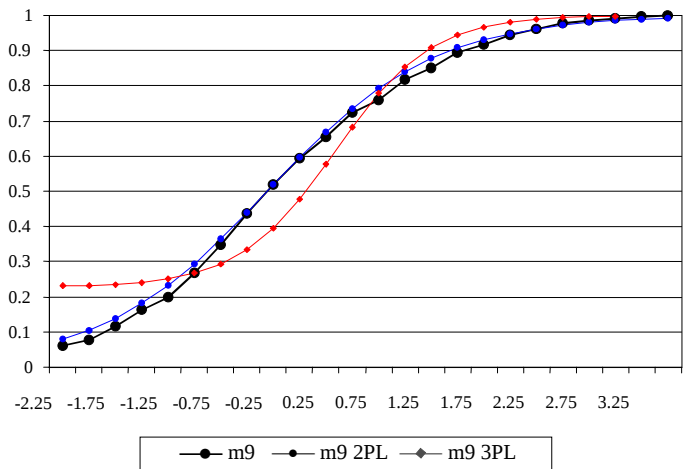
TABLA 3: PESOS FACTORIALES DE LOS 48 ÍTEMES DE LA PRUEBA SIMCE DE MATEMÁTICAS

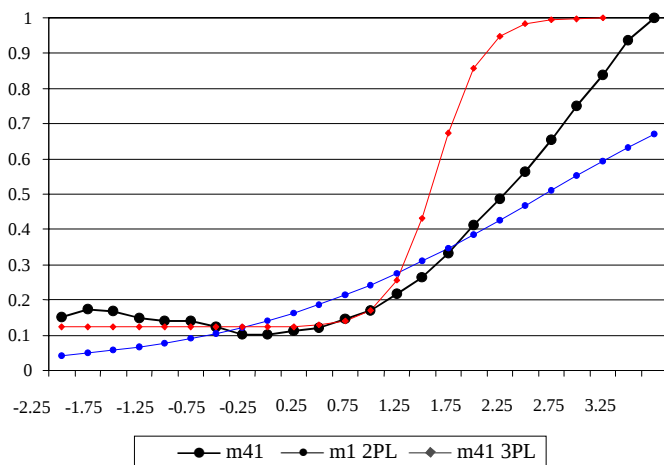
Ítem	Peso factorial	Ítem	Peso factorial
m01	0.44	m25	0.03
m02	0.50	m26	0.34
m03	0.40	m27	0.38
m04	0.52	m28	0.32
m05	0.16	m29	0.58
m06	0.50	m30	0.57
m07	0.32	m31	0.48
m08	0.54	m32	0.59
m09	0.55	m33	0.50
m10	0.39	m34	0.60
m11	0.51	m35	0.58
m12	0.38	m36	0.62
m13	0.31	m37	0.45
m14	0.54	m38	0.69
m15	0.54	m39	0.49
m16	0.44	m40	0.32
m17	0.47	m41	0.38
m18	0.53	m42	0.48
m19	0.52	m43	0.41
m20	0.43	m44	0.52
m21	0.59	m45	0.49
m22	0.56	m46	0.33
m23	0.66	m47	0.24
m24	0.51	m48	0.56

6.2. Comparación del ajuste de distintos modelos IRT

Tal como fue previamente mencionado, hay modelos IRT de 1, 2 y 3 parámetros. Una de las decisiones básicas en aplicaciones prácticas de este enfoque se refiere a la adopción de una de estas alternativas. Desde el punto de vista conceptual, el modelo de 3 parámetros sólo es pertinente cuando se está trabajando con preguntas de respuesta cerrada, donde existe la posibilidad de beneficiarse contestando al azar. Excepto por este criterio, que definitivamente excluye el modelo de 3 parámetros cuando no hay probabilidad de responder correctamente por azar, la opción entre los modelos debiera fundarse en un examen del grado en que cada uno de los modelos se ajusta a los datos. Aunque existen algunas aproximaciones estadísticas para apoyar esta decisión, en la práctica, siempre se recomienda el examen empírico del grado en que las curvas características resultantes de la aplicación de modelos de diverso número de parámetros, concuerdan con las curvas empíricas correspondientes. Las curvas empíricas se obtienen por un procedimiento análogo al empleado para construir las curvas descritas en la Figura 1, sólo que en este caso se usa como estimación de la habilidad de los examinados la que determina el modelo IRT (θ). En el caso de las pruebas SIMCE de matemática, se comparó el modelo de 2 parámetros con el de 3 parámetros. Dado que se trata de preguntas de selección múltiple, resultaba lógico contemplar la posibilidad que el modelo de 3 parámetros obtuviera un mejor ajuste que uno de 2 parámetros. Sin embargo, la mayoría de los ítems de esta prueba mostró un mejor grado de ajuste del modelo de 2 parámetros. Los gráficos de la Figura 4 ilustran, con fines descriptivos,

FIGURA 4: COMPARACIÓN ENTRE CURVAS EMPÍRICAS Y CCI DE 2 Y 3 PARÁMETROS EN 2 ÍTEMES DE LA PRUEBA SIMCE DE MATEMÁTICA



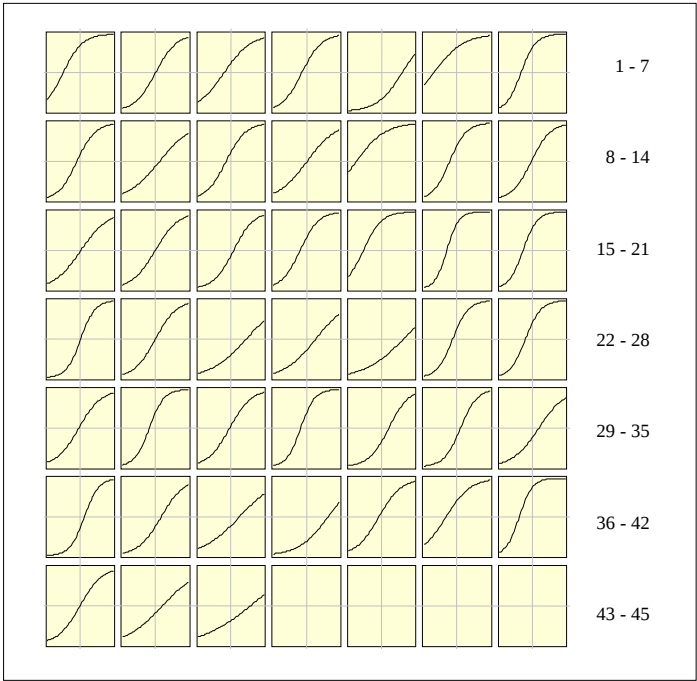


sólo dos ítemes: uno que muestra un buen ajuste al modelo de 2 parámetros, y otro con un pobre ajuste. Este último ítem corresponde al tercer ítem excluido de esta prueba.

6.3. Calibración de modelos IRT

La calibración de los ítemes al modelo IRT corresponde a la fase de estimación de los parámetros en el modelo correspondiente. Para llevar a cabo esta tarea se requiere el empleo de software especializado. Aunque se han desarrollado varios programas con este propósito, los más populares son los producidos por Darrell Bock y sus colaboradores: BILOG, PARSCALE y MULTILOG. En conjunto, estos programas permiten resolver una amplia variedad de problemas prácticos, incluyendo modelos de diverso número de parámetros, para ítemes puntuados dicotómica o policotómicamente. En este último caso, los programas ofrecen diversas opciones, como el modelo de respuesta graduada de Samejima, el de crédito parcial de Muraki o el de respuesta nominal de Bock, entre otros. Asimismo, ofrecen diversos algoritmos para la estimación de los parámetros o las habilidades de los examinados. Por último, cuentan con opciones para manejar múltiples grupos (lo que es conveniente para realizar análisis DIF), así como múltiples examinadores (cuando diversos jueces deben asignar puntajes a respuestas abiertas). Las versiones más recientes de estos programas incluyen una interfaz gráfica que permite visualizar diversos aspectos, tales como las curvas características de los ítemes, su función informativa, la función informativa del test, etc.

FIGURA 6: CURVAS CARACTERÍSTICAS DE LOS ÍTEMES DE LA PRUEBA SIMCE DE MATEMÁTICA



En el caso que nos interesa, la calibración final de la prueba se llevó a cabo con PARSCALE 4.1 (Muraki y Bock, 2003) empleando el modelo de 2 parámetros. Luego de 8 ciclos se obtuvo una solución satisfactoria: los parámetros resultantes se encuentran en un rango adecuado, con un valor medio de 0,61 para el parámetro de discriminación, y $-0,06$ para el parámetro de localización o dificultad. Además, tal como puede apreciarse en la Tabla 1, el error estándar asociado a todos los parámetros es homogéneamente bajo, lo que respalda el ajuste de los ítems al modelo. La Figura 5, por último, muestra gráficamente las curvas características de los 45 ítems de esta prueba.

6.4. Evaluación de la capacidad informativa de una prueba

La función informativa de la prueba es un importante indicador de la calidad de un instrumento. Esta función nos revela el grado de precisión de un instrumento para evaluar el atributo subyacente en distintas regiones del mismo. Como es sabido, en la Teoría Clásica de la Medición el concepto que revela la calidad de la medición, la confiabilidad, consiste en un indica-

do único, que alude a la calidad global de un instrumento. Lamentablemente, es engañoso creer, como se mencionó en la sección 5.2., que un instrumento pueda ser homogéneamente preciso. El error de medición no es constante, por lo que es conveniente contar con una evaluación más específica y precisa de la calidad de la medición.

En el caso de la prueba que estamos analizando, su función informativa, que aparece en la Figura 1, revela que el instrumento es adecuadamente informativo en la región media de los puntajes. Dicha precisión decae en ambos extremos. Dado que el propósito de este instrumento es estimar la habilidad de grupos de estudiantes (pues el SIMCE se reporta a nivel de establecimientos), se puede constatar que para la gran mayoría de los puntajes que se estiman el instrumento produce una información adecuada.

6.5. Análisis del sesgo de preguntas (análisis DIF)

El sesgo de medición se ha transformado, en los últimos años, en un tema de gran importancia en la medición. En la medida que los instrumentos de medición han adquirido gran relevancia para tomar importantes decisiones, con claras consecuencias personales o sociales¹³, la preocupación por la posibilidad que los instrumentos puedan inadvertidamente favorecer o perjudicar a determinados grupos, ha aumentado considerablemente. Ello ha repercutido directamente en la teoría de la medición, reflejándose en el desarrollo de técnicas y procedimientos para evaluar el sesgo en la medición, tanto a nivel de ítemes como de los puntajes totales de un instrumento. Los estándares vigentes para el desarrollo de instrumentos que tienen consecuencias sociales y personales relevantes establecen la necesidad de evaluar dicho sesgo.

El análisis del funcionamiento diferencial de los ítemes (conocido por su sigla en inglés DIF), se ha transformado en el enfoque dominante para estos propósitos. Con esta técnica se busca establecer si un determinado grupo (establecido a partir de diferencias de género, raciales, sociales u otras) puede verse beneficiado o perjudicado en sus puntajes (Camilli y Shepard, 1994). El enfoque IRT ofrece un marco de referencia muy claro para esta evaluación, puesto que, de acuerdo a sus fundamentos, la posibilidad de responder correctamente un ítem debiera depender sólo de la habilidad de una persona. En consecuencia, cuando la pertenencia a un grupo

¹³ Para ilustrar esto basta pensar en el rol que hoy tienen en todo el mundo, incluido nuestro país, instrumentos de medición en decisiones como: acceso a la educación (primaria, secundaria y universitaria), acceso y evaluación laboral (selección laboral, evaluación del desempeño, etc.), peritajes judiciales, etc.

afecta los puntajes independientemente de su nivel de habilidad, se produce una violación de los supuestos fundamentales de la teoría. Se dice que un ítem tiene un efecto DIF cuando la pertenencia grupal de un examinado afecta su probabilidad de responder correctamente dicho ítem, más allá de su nivel de habilidad.

Para ilustrar este tipo de análisis emplearemos nuevamente información proveniente de las mediciones SIMCE. Sin embargo, deberemos recurrir a una medición distinta de la empleada en los ejemplos anteriores, puesto que para facilitar la interpretación del efecto DIF es conveniente presentar el contenido de los ítems involucrados. Dado que los ítems empleados en la medición SIMCE de 2º medio de 1998 fueron liberados al conocimiento público, podemos recurrir a ellos. En concreto, se presentan a continuación 2 ítems de la prueba de lenguaje, uno con un efecto DIF favorable a las mujeres y otro con un efecto DIF favorable a los hombres. En ambos casos se trata de preguntas que se basan en un texto previo (omitido aquí, pues no resulta indispensable para interpretar los resultados).

Ejemplo 1: Ítem con efecto DIF favorable a las mujeres

¿Cuál es el principal sentimiento que expresa el relator hacia la mujer descrita?

- a) Simpatía
- b) Curiosidad
- c) Amor
- d) Crítica

Ejemplo 2: Ítem con efecto DIF favorable a los hombres

En el texto, ¿qué significa “Europa fue el teatro inicial del conflicto”?

- a) Que la Segunda Guerra Mundial comenzó en Europa
- b) Que al principio Europa fue una espectadora del conflicto
- c) Que Europa comenzó la Segunda Guerra Mundial
- d) Que al principio la Segunda Guerra Mundial fue una farsa

Para evaluar el efecto DIF se empleó el software PARSCALE, que ofrece la posibilidad de estimar modelos IRT con múltiples grupos. En concreto, se evaluó un modelo DIF comparando a hombres y mujeres. El análisis DIF se centró en el parámetro de localización o dificultad. Los resultados, que se representan gráficamente en la Figura 6a y 6b, muestran que en estos dos ítems el parámetro b mostró diferencias estadísticamente

significativas. En el primer caso, la curva correspondiente a los hombres aparece desplazada hacia la derecha, lo que indica que en este ítem es necesario un mayor nivel de habilidad para responderlo correctamente. Lo contrario ocurre con el segundo ítem. La interpretación del efecto DIF en estos dos casos resulta relativamente simple, puesto que el contenido de ambas preguntas se vincula de manera más o menos directa con los aspectos tradicionalmente asociados a la socialización diferencial de hombres y mujeres: la primera pregunta alude a un tema usualmente considerado más femenino, la comprensión de emociones, mientras que la segunda se relaciona con un tema convencionalmente visto como masculino, la guerra.

Es importante dejar constancia que la detección de un efecto DIF no debe ser considerada como una razón para eliminar automáticamente un ítem. Más bien, esta evidencia debe ser incorporada como una señal de alerta, que debe ser analizada en conjunto con otros antecedentes para resolver la suerte de un ítem. En el caso que nos ocupa, el efecto, que resulta fácilmente interpretable, nos alerta acerca de los riesgos de incluir preguntas cuyo contenido o temática pueda resultar más familiar, cercana o significativa para un determinado grupo.

FIGURA 6a: CURVAS CARACTERÍSTICAS DE ÍTEM CON EFECTO DIF FAVORABLE A MUJERES

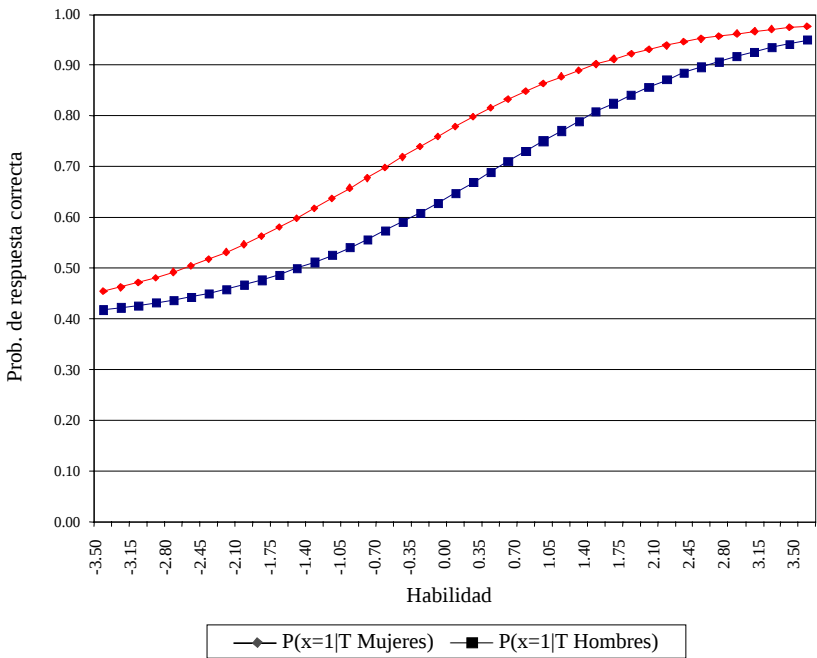
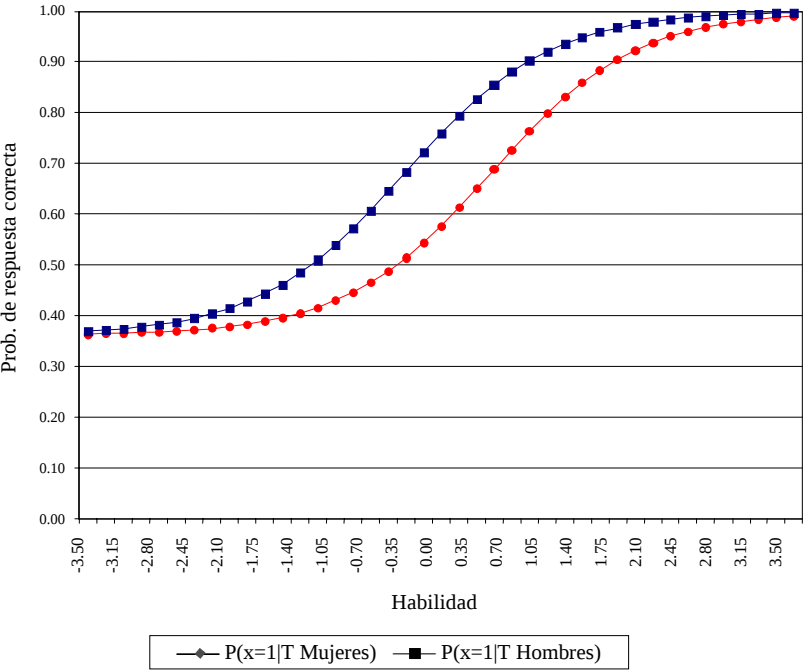


FIGURA 6b: CURVAS CARACTERÍSTICAS DE ÍTEM CON EFECTO DIF FAVORABLE A HOMBRES



Conclusiones

En este trabajo hemos revisado los fundamentos de las dos principales teorías de la medición desarrolladas durante el siglo pasado: la Teoría Clásica y la teoría moderna o Teoría de Respuesta al Ítem. Hemos hecho ver que ambas teorías fundan sus postulados básicos en supuestos semejantes, aunque difieren en el grado de complejidad con respecto a la forma en que modelan la relación entre los atributos subyacentes y las variables manifiestas (respuestas a las preguntas de un instrumento de medición). En gran medida el modelo IRT es una extensión de la TCM, por lo que estas teorías no pueden ser vistas como modelos competitivos o antagonicos.

Así como la TCM, los postulados básicos del enfoque IRT, al menos en lo referente al análisis de instrumentos de medición unidimensionales con preguntas puntuadas dicotómicamente, constituyen un cuerpo asentado

y bien establecido de principios, por lo que su empleo para resolver problemas prácticos de medición en el ámbito psicológico y educacional se encuentra sólidamente respaldado.

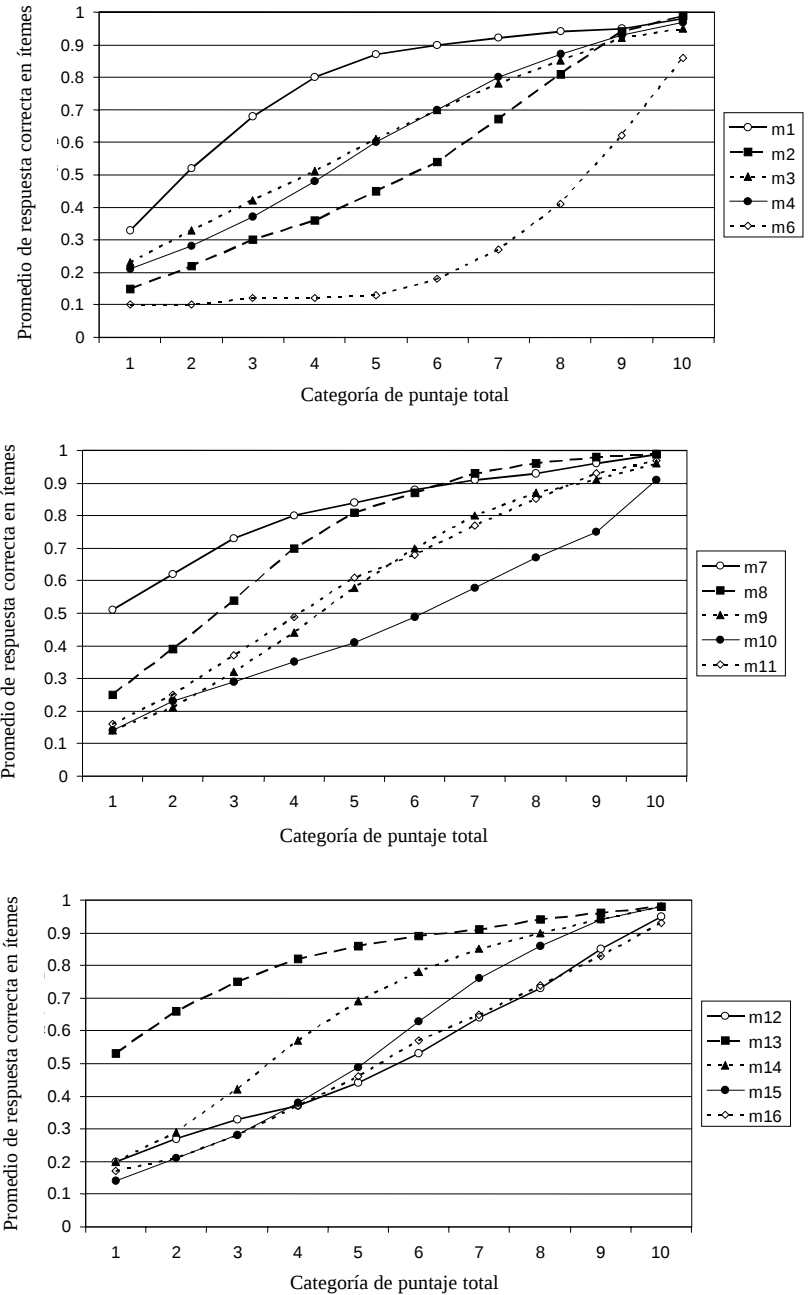
Es cierto que el modelo IRT ha demorado en ser incorporado masivamente en el desarrollo de instrumentos de medición. Excepto por las aplicaciones a la medición educacional en gran escala, su implementación en otros campos ha sido relativamente reciente, tal como se ilustró en la introducción de este trabajo. Esta demora, sin embargo, no refleja debilidades de la teoría, sino que barreras de orden práctico: la relativa escasez de programas computacionales capaces de manejar en forma eficiente y rápida los complejos algoritmos asociados a la estimación de los parámetros de los ítems, las exigencias muestrales relativamente altas de esta teoría, relativas a las existentes en la TCM, y la mayor complejidad matemática de este enfoque en relación a su precedente.

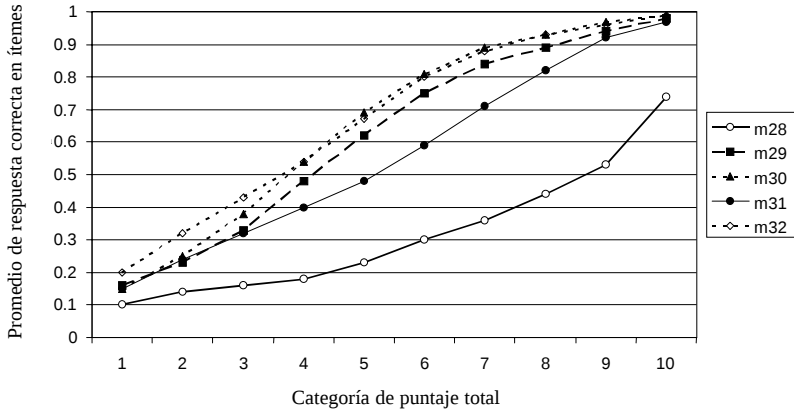
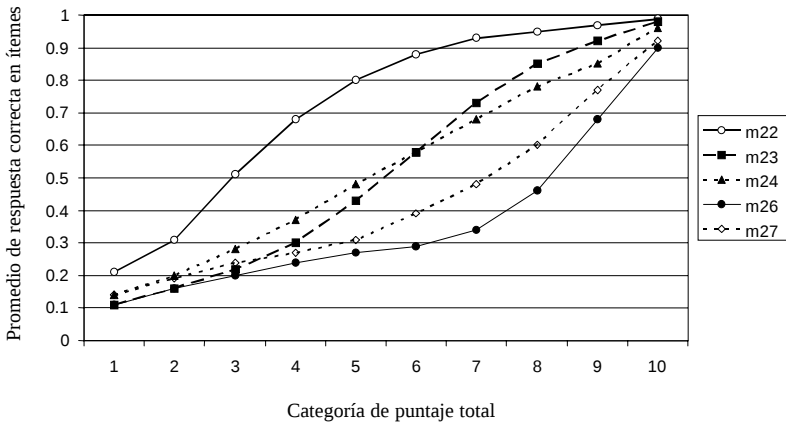
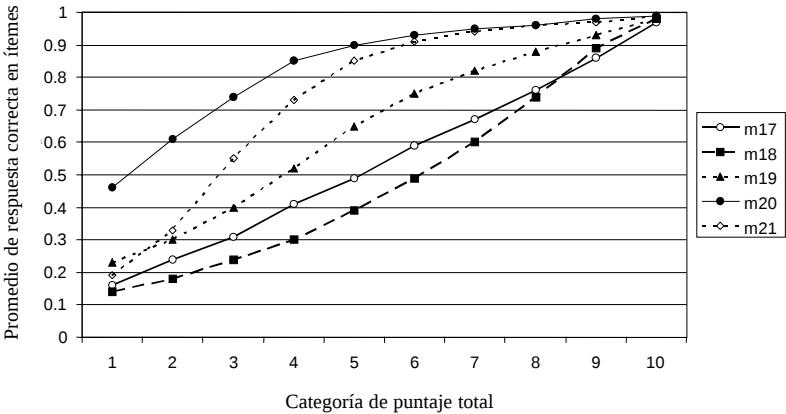
Es importante destacar que el interés académico en torno a la teoría de la medición se ha volcado casi enteramente en los últimos años a la Teoría de Respuesta al Ítem. Basta revisar el índice de las principales revistas científicas relacionadas con la psicometría (*Psychometrika*, *Applied Psychological Measurement* y *Psychological Methods*) para constatar que casi todos los artículos que se publican se refieren a este enfoque. De hecho, un examen de la revista más especializada en teoría de la medición, *Applied Psychological Measurement*, muestra que de los 25 artículos publicados en 2002, 24 tenían relación con IRT y 1 con la teoría de la generalizabilidad (ninguno se refería a la teoría clásica). Este dinamismo académico, sumado al creciente desarrollo de software especializado, permite anticipar que en los próximos años se verá un aumento sostenido del empleo de métodos basados en este enfoque para diseñar y evaluar instrumentos de medición tanto en educación como en psicología.

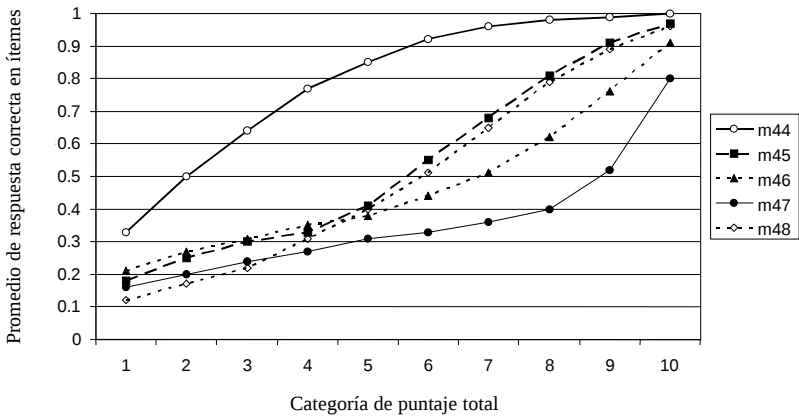
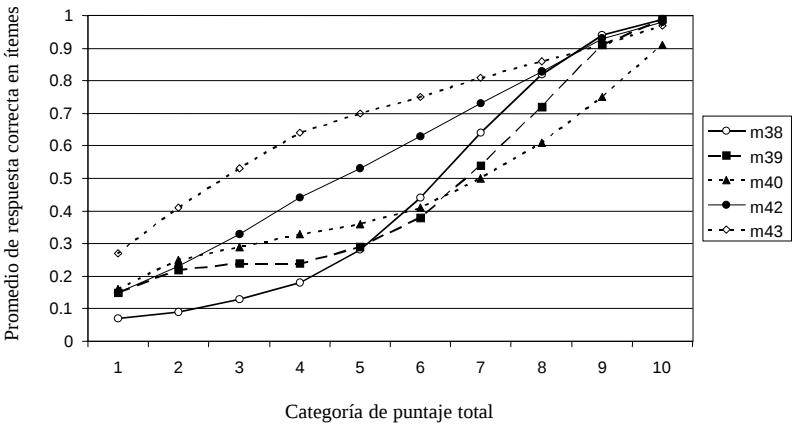
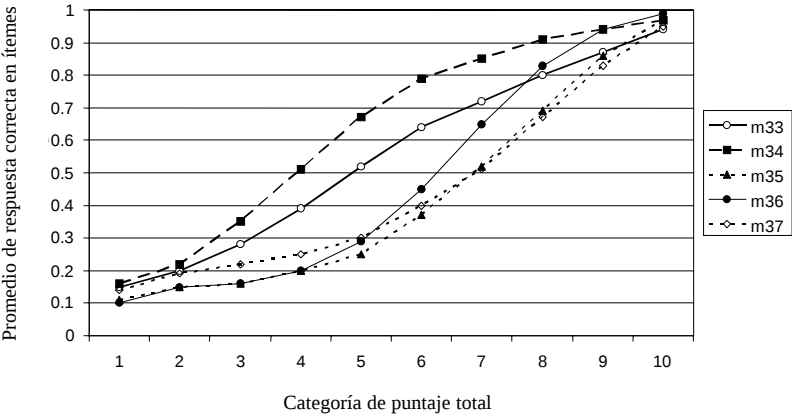
En consecuencia, lo único razonable es incorporar este enfoque en la agenda de trabajo de todo esfuerzo serio en materia de medición. No se trata, como esperamos haber clarificado en este artículo, de reemplazar a la Teoría Clásica, sino de complementarla con los procedimientos, técnicas y posibilidades que ofrece el enfoque IRT. Especialmente en los casos en que se dispone de grandes volúmenes de información, como es el caso de la medición educacional en gran escala, no hay razones científicamente fundadas para excluir el uso de este enfoque.

ANEXO

FIGURA 1: CURVAS EMPÍRICAS DE LOS ÍTEMES DE LA PRUEBA SIMCE DE MATEMÁTICA







BIBLIOGRAFÍA

- Andersen, E.B., y Olsen, L. W. "The Life of George Rasch as a Mathematician and as a Statistician". En A. Boomsma *et al.* (eds.), *Essays on Item Response Theory*. New York: Springer, 2001.
- Anderson, T. W. "Some Scaling Models and Estimation Procedures in the Latent Class Model". En U. Grenander (ed.), *Probability and Statistics*. New York: Wiley, 1959, pp. 9-38.
- Bartholomew, D. J. *Latent Variable Models and Factor Analysis*. Londres: Charles Griffin, 1987.
- Birnbaum, A. "On the Foundations of Statistical Inference (with Discussion)". *Journal of the American Statistical Association* 57 (1962), pp. 269-326.
- Bock, D., Gibbons, R., Schilling, S., Muraki, E., Wilson, D., y Wood, R. TESTFACT 4.0. Test scoring, item statistics, and item factor analysis [Programa computacional]. Chicago: Scientific Software International, 2003.
- Camilli, G., y Shepard, L. A. *Methods for Identifying Biased Test Items*. Thousand Oaks, CA: Sage, 1994.
- Cantor, G. "Beiträge Zur Begründung der Transfiniten Mengenlehre". *Math. Ann.* 46 (1895), pp. 481-512.
- Cox, D. R. "Role of Models in Statistical Analysis". *Statistical Sciences* 5 (1990), pp. 169-174.
- Crocker, L., y Algina, J. *Introduction to Classical and Modern Test Theory*. Philadelphia: Harcourt Brace Jovanovich College Publishers, 1986.
- Cronbach, L., Gleser, G., y Rajaratnam, N. "Theory of Generalizability. A Liberation of Reliability Theory". *British Journal of Mathematical and Statistical Psychology*, 16 (1963), pp. 137-173.
- Dusaillant, F. "Técnicas de Medición en Pruebas de Admisión a las Universidades". *Estudios Públicos* 90 (otoño 2003).
- Ellis, B. *Basic Concepts of Measurement*. Cambridge: Cambridge University Press, 1968.
- Ellis, B. "Differential Item Functioning: Implications for Test Translations". *Journal of Applied Psychology*, 74 (1989), pp. 912-921.
- Embretson, S. E., y Reise, S. P. *Item Response Theory for Psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates, 2000.
- Fisher, R. A. "On the Mathematical Foundations of Theoretical Statistics". *Philosophical Transaction of the Royal Society of London, Series A* 222 (1922), pp. 309-368.
- Fischer, G. H. "Applying the Principles of Specific Objectivity and Generalizability to the Measurement of Change". *Psychometrika* 52 (1987), pp. 565-587.
- Fischer, G. H. "Some Neglected Problems in IRT". *Psychometrika* 60 (1995a), pp. 459-487.
- Fischer, G. H. "Derivations of the Rasch Mode". En G. H. Fischer e I. W. Molenaar (eds.), *Rasch Models. Foundations, Recent Developments and Applications*. New York: Springer, 1995b.
- Fraley, R., Waller, N., y Brennan, K. "An Item Response Theory Analysis of Self-Report Measures of Adult Attachment". *Journal of Personality and Social Psychology*, 78 (2000), pp. 350-365.
- Goldberger, A. S. "Econometrics and Psychometrics: A Survey of Communalities". *Psychometrika* 36 (1971), pp. 83-107.
- Goldberger, A. S. "Structural Equation Methods in the Social Sciences". *Econometrica* 40 (1972), 979-1001.

- Godber, T., Anderson, V., y Bell, R. "The Measurement and Diagnostic Utility of Intrasubtest Scatter in Pediatric Neuropsychology". *Journal of Clinical Psychology*, 56 (2000), pp. 101-112.
- Gray-Little, B., Williams, V., y Hancock, T. "An Item Response Theory Analysis of the Rosenberg Self-Esteem Scale". *Personality and Social Psychology Bulletin*, 23 (1997), pp. 443-451.
- Gulliksen, H. *Theory of Mental Tests*. New York: Wiley, 1950.
- Guttman, L. "A Basis for Analyzing Test-Retest Reliability". *Psychometrika* 10 (1945), pp. 255-273.
- Guttman, L. "Reliability Formulas that do not Assume Experimental Independence". *Psychometrika* 18 (1953), pp. 225-239.
- Hambleton, R. K., Swaminathan, H., y Rogers, H. J. *Fundamentals of Item Response Theory*. Newbury Park, CA: Sage, 1991.
- Hand, D. J. "Statistics and the Theory of Measurement (with Discussion)". *Journal of the Royal Statistical Society, Series A* 159 (1996), pp. 445-492.
- Helmholtz, H. V. "Zählen und Messenerkenntnis-Theoretisch Betrachtet". *Philosophische Aufsätze Eduard Zeller gewidmet*, Leipzig, 1895.
- Holland, P. W., y Rosenbaum, P. R. "Conditional Association and Unidimensionality in Monotone Latent Variable Models". *The Annals of Statistics* 14 (1980), pp. 1523-1543.
- Hölder, O. "Die Axiome der Quantität und Die Lehre von Mass". *Ber. Verh. Kgl. Sächsis. Ges. Wiss. Leipzig, Math.-Phys. Classe* 53 (1901), pp. 1-64.
- Junker, B. "On the Interplay Between Nonparametric and Parametric IRT, with some Thoughts about the Future". En A. Boomsma *et al.* (eds.), *Essays on Item Response Theory*. New York: Springer, 2001.
- Koopmans, T. C., y Reiersøl, O. "Identification of Structural Characteristics". *The Annals of Mathematical Statistics* 21 (1950), pp. 165-181.
- Krantz, D. H., Luce, R. D., Suppes, P., y Tversky, A. *Foundations of Measurement. Volume I. Additive and Polynomial Representations*. New York: Academic Press, 1971.
- Lazarsfeld, P. F. "The Logical and Mathematical Foundation of Latent Structure Analysis". En S. A. Stouffer *et al.* (eds.), *Measurement and Prediction*. New York: Wiley, 1959.
- Lindley, D. V., y Novick, M. R. "The Role of Exchangeability in Inference". *The Annals of Statistics* 9 (1981), pp. 45-58.
- Lord, F. M., y Novick, M. R. *Statistical Theories of Mental Test Scores*. Massachusetts: Addison Wesley, 1968.
- Maller, S. "Item Invariance in Four Subtests of the Universal Nonverbal Intelligence Test (UNIT) Across Groups of Deaf and Hearing Children". *Journal of Psychoeducational Assessment*, 18 (2000), pp. 240-254.
- Manski, C. F. *Identification Problems in the Social Sciences*. Cambridge, MA: Harvard University Press, 1995.
- McCullagh, P. "What is a Statistical Model? (with Discussion)". *The Annals of Statistics* 30 (2002), pp. 1225-1310.
- Michell, J. *An Introduction to the Logic of Psychological Measurement*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1990.
- Michell, J. "Bertrand Russell's 1897 Critique of the Traditional Theory of Measurement". *Synthese* 110 (1997a), pp. 257-276.

- Michell, J. "Quantitative Science and the Definition of Measurement in Psychology (with Discussion)". *British Journal of Psychology* 88 (1997b), pp. 355-383.
- Mislevy, R. J. "Recent Developments in the Factor Analysis of Categorical Variables". *Journal of Educational Statistics*, 11 (1986), pp. 3-31.
- Mouchart, M., y San Martín, E. "Specification and Identification Issues in Models Involving a Latent Hierarchical Structures". *Journal of Statistical Planning and Inference* 111 (2003), pp. 143-163.
- Muraki, E., y Bock, D. PARSCALE 4.1: IRT item analysis and test scoring for rating-scale data [Programa computacional]. Chicago: Scientific Software International, 2003.
- Muthén, L., y Muthén, B. *Mplus: Statistical Analysis with Latent Variables* [programa computacional]. Los Angeles: Muthen y Muthen, 1998.
- Novick, M. R., y Lewis, C. "Coefficient Alpha and the Reliability of Composite Measurements". *Psychometrika* 32 (1967), pp. 1-13.
- Novick, M. R. "The Axioms and Principal Results of Classical test Theory". *Journal of Mathematical Psychology* 3 (1968), pp. 1-18.
- Novick, M. R. "Statistics as Psychometrics". *Psychometrika* 45 (1980), pp. 411-424.
- Panter, A., Swygert, K., Dahlstrom, W., y Tanaka, J. "Factor Analytic Approaches to Personality Item-Level Data". *Journal of Personality Assessment*, 68 (1997), 561-589.
- Pfanzagl, J. *Theory of Measurement*. Physica-Verlag, Würzburg, 1968.
- Rouse, S., Finger, M., y Butcher, J. "Advances in Clinical Personality Measurement: An Item Response Theory Analysis of the MMPI-2 PSY-5 Scales". *Journal of Personality Assessment*, 72 (1999), pp. 282-307.
- San Martín, E. "Modeling Problems Motivated by the Specification of Latent Linear Structures". Aceptado para publicación en el *Journal of Mathematical Psychology*, 2003.
- Sobel, M. E. "Measurement, Causation and Local Independence in Latent Variable Models". En M. Berkane (ed.), *Latent Variable Modeling and Applications to Causality*. New York: Springer, 1997.
- Swistak, P. "Paradigms of Measurement". *Theory and Decision* 29 (1990), pp. 1-17.
- Thissen, D., y Wainer, H. *Test Scoring*. Mahwah, NJ: Lawrence Erlbaum Association, 2001.
- Van der Linden, W., y Hambleton, R. *Handbook of Modern Item Response Theory*. New York: Springer, 1997.
- Waller, N., Thompson, J., y Wenk, E. "Using IRT to Separate Measurement Bias from True Group Differences on Homogeneous and Heterogeneous Scales: An Illustration with the MMPI". *Psychological Methods*, 5 (2000), pp. 125-146.
- Woods, C. M. "Factor Analysis of Scales Composed of Binary Items: Illustration with the Maudsley Obsessional Compulsive Inventory". *Journal of Psychopathology and Behavioral Assessment*, 24 (2002), pp. 215-223. □

**TEORÍA CLÁSICA DE MEDICIÓN
O TEORÍA DE RESPUESTA AL ÍTEM**
LA EXPERIENCIA SUECA

Christina Stage

En este trabajo se intenta determinar si la aplicación de la teoría de respuesta al ítem (IRT, por sus siglas en inglés) permitiría mejorar la calidad de la prueba de admisión a las universidades en Suecia (SweSAT). Conforme a la evidencia analizada, la autora concluye que el uso de IRT no proporciona ninguna ventaja por sobre el método clásico en una prueba como la SweSAT. Con todo, señala que la teoría de respuesta al ítem está aún en desarrollo y que, por cierto, hay áreas específicas en que ella pudiera resultar muy promisoria. Por último, la autora destaca el caso de los tests adaptativos computarizados (TAC) o “tests a la medida”, donde IRT no sólo proporciona un marco teórico adecuado, sino el único factible en esas circunstancias. Si en el futuro la SweSAT —o alguna versión de la misma— llegara a transformarse en un TAC, entonces forzosamente tendría que usarse IRT. Pero de mantenerse la SweSAT en su esquema actual, no se recomienda el uso de IRT en su confección.

CHRISTINA STAGE. Profesora del Departamento de Medición Educacional, Universidad de Umeå, SE 901 87 Umeå, Suecia. christina.stage@edmeas.umu.se

Este trabajo describe los esfuerzos realizados para dilucidar si la teoría de respuesta al ítem (IRT por sus siglas en inglés) sería aplicable a la Prueba de Admisión a las Universidades en Suecia (SweSAT). El objetivo ha sido determinar si un cambio desde la teoría clásica de la medición (TCM) hacia la IRT en el proceso de creación de ítems, diseño de pruebas, asignación de puntajes o “equating”^{*} permitiría mejorar la calidad de la prueba. El trabajo consta de tres partes. La primera parte, “La aplicabilidad de modelos de IRT a los subtests de SweSAT”, es un resumen de cinco informes (Stage, 1996, 1997, 1997b, 1997c y 1997d) que describen las diferentes etapas que se han seguido para investigar si un modelo de IRT podría ser ajustado a los cinco subtests de SweSAT por separado. La segunda parte, “Comparación entre análisis de ítems basándose en IRT y TCM”, contiene una síntesis de tres informes anteriores (Stage, 1988a, 1988b y 1999), en los cuales se efectuaron comparaciones entre índices de dificultad y discriminación dentro de la TCM y parámetros de dificultad y discriminación dentro de la IRT en los tres subtests ERC, READ y WORD. En la tercera parte, “Aplicabilidad de la IRT al SweSAT: el test total”, se describe un intento por ajustar un modelo de IRT a la SweSAT total. La conclusión fue que como el ajuste del modelo a los datos resultó algo dudoso, especialmente para el test total, no se obtenía ninguna ventaja al cambiar de TCM a IRT.

La Prueba de Admisión a las Universidades en Suecia (SweSAT) es un test referido a normas, que se emplea en el proceso de selección para ingresar a la educación superior en Suecia. Se aplica dos veces al año, una durante el trimestre de primavera y otra durante el trimestre de otoño. Luego de ser rendida, cada prueba en particular pasa a ser del dominio público, de modo que es preciso elaborar una nueva versión de ésta cada vez que se vuelve a administrar. Puesto que los resultados de la prueba tienen una validez de 5 años, es importante que los resultados de diferentes aplicaciones de la prueba sean comparables.

Desde 1996 la prueba está constituida por 122 ítems de selección múltiple, divididos en 5 subtests:

1. DS: Subtest de suficiencia de datos que mide razonamiento matemático y consta de 22 ítems.
2. DTM: Subtest que mide la habilidad para interpretar diagramas, tablas y gráficos, que consta de 20 ítems.

^{*} Proceso por el cual los puntajes de diferentes pruebas se hacen comparables. (N. del E.)

3. ERC: Subtest de comprensión de lectura en inglés que consta de 20 ítems.
4. READ: Subtest de comprensión de lectura en sueco, que consta de 20 ítems.
5. WORD: Subtest de vocabulario, que consta de 40 ítems.

Desde que se aplicó la primera versión de SweSAT en 1977, la elaboración y la compilación del test, lo mismo que el “equating” de las formas de éste entre una administración y la próxima, han estado basados en la teoría clásica de medición (TCM).

En TCM, que comenzó a evolucionar con el test de Binet hace casi cien años, se considera que el puntaje del test está constituido por dos elementos, un “puntaje real” y un error. El puntaje real y el error aparecen como factores completamente independientes. El puntaje real es visto como un componente que no varía entre una forma de test y una forma paralela alternativa, y entre una y otra ocasión. Se considera que el error es una característica única de la medición específica, y es enteramente independiente del error que cabría esperar que surgiera en otra medición del mismo constructo. El puntaje real nunca se puede observar directamente, y sólo puede inferirse a partir de la coherencia en el desempeño teniendo en cuenta el puntaje obtenido de un test a otro.

La TCM ha sido un modelo productivo que condujo a la formulación de una serie de relaciones útiles:

- La relación entre la longitud del test y su precisión (confiabilidad).
- Estimaciones de la precisión de diferencias de puntajes y cambios de puntajes.
- La estimación de las propiedades de compuestos de dos o más medidas.
- La estimación del grado en que los índices de relación entre distintas mediciones son atenuados por el error de medición en cada una de ellas.

Aun cuando el principal centro de interés de la TCM es la información a nivel del test, las estadísticas de los ítems (es decir el nivel de dificultad de los ítems y el nivel de discriminación de los mismos) también son importantes. A nivel de ítems la TCM es relativamente sencilla, ya que no existen modelos teóricos complejos para relacionar la habilidad o el éxito de un examinado en un ítem específico. La proporción de un grupo claramente definido de examinados que responden un ítem correctamente (determinada de manera empírica) —valor p — se emplea como índice de

dificultad del ítem (en verdad se trata de un indicador inverso de dificultad, pues los valores más altos señalan que los ítems son más fáciles). La capacidad de un ítem para discriminar entre examinados de alta habilidad y de baja habilidad se expresa estadísticamente como el coeficiente de correlación entre los puntajes logrados en el ítem y los puntajes obtenidos en el test total.

Con frecuencia se alude a los modelos de TCM como modelos “débiles”, porque sus supuestos son fácilmente confirmados en los datos del test.

La TCM adolece, sin embargo, de algunas deficiencias. Una de ellas es que los índices de dificultad y de discriminación de los ítems son dependientes del grupo; en efecto, los valores de estos índices dependen de los grupos de examinados en que se han obtenido. Otra limitación es la dependencia entre los puntajes observados y reales de los tests. Los puntajes observados y reales suben y bajan de acuerdo con los cambios en la dificultad de los tests. Otro punto débil tiene que ver con el supuesto de que los errores de medición son los mismos para todos los examinados. Las estimaciones de habilidad son, en realidad, menos precisas para alumnos de baja y alta habilidad que para los alumnos de habilidad promedio.

Durante las últimas décadas se ha desarrollado un nuevo sistema de medición, llamado teoría de respuesta al ítem (IRT por sus siglas en inglés), que ha llegado a transformarse en un importante complemento de la TCM en el diseño, la construcción y la evaluación de pruebas. Dentro del marco de IRT es posible obtener características de los ítems que *no* dependen de grupos; puntajes de habilidad que *no* dependen de los tests; y una medición de la precisión para cada nivel de habilidad.

De acuerdo con Hambleton y otros (1991):

IRT se funda en dos postulados básicos: a) el desempeño de un examinado en el ítem de un test puede predecirse (o explicarse) mediante una serie de factores llamados rasgos, rasgos latentes o habilidades; y b) la relación entre el desempeño de un examinado en un ítem y la serie de rasgos implícitos en el desempeño en el ítem puede describirse con una función que aumenta monótonamente, llamada función característica del ítem o curva característica del ítem. Esta función especifica que a medida que el nivel del rasgo aumenta, la probabilidad de responder correctamente también aumenta (p. 7).

Existen diversos modelos de IRT, pero todos tienen en común el uso de una función matemática para especificar la relación entre el desempeño observable del examinado en un test y los rasgos o habilidades no observables que se supone están implícitos en el desempeño en el test. En cualquier

aplicación práctica de los modelos de rasgos latentes es preciso especificar la forma matemática de las curvas características del ítem y obtener estimaciones de los parámetros del ítem necesarios para describir las curvas. En el modelo de tres parámetros éstos son:

1. Dificultad del ítem “b”.
2. Discriminación del ítem “a”.
3. Un parámetro de pseudoadivinanza “c”.

En el modelo de dos parámetros no se supone que exista alguna adivinanza, mientras que en el modelo de un parámetro se supone que la discriminación del ítem es la misma para todos los ítems.

Los modelos de IRT se denominan modelos “fuertes”, ya que los supuestos pueden resultar difíciles de confirmar por los datos del test. Un importante supuesto incluido en los modelos más comunes de IRT es el de la unidimensionalidad, que significa que los ítems que constituyen el test sólo miden una habilidad. Lo que se requiere para que el supuesto de unidimensionalidad se confirme adecuadamente es la presencia de un factor predominante que influya en el desempeño en el test. Otro supuesto, relacionado con el anterior, es el de la independencia local, la cual indica que cuando las habilidades que influyen en el desempeño en el test se mantienen constantes, las respuestas del examinado a cualquier par de ítems son estadísticamente independientes.

Una vez que se especifica un modelo de rasgos latentes, la precisión con que éste estima la habilidad del examinado se puede determinar para distintos niveles de habilidad. La información varía con el nivel de habilidad, lo que hace posible determinar el error estándar de estimación para distintos niveles de habilidad. La función de información del ítem entrega información sobre la utilidad del ítem para medir la habilidad en un determinado nivel.

En la actualidad, las entidades que elaboran tests están prestando creciente atención a IRT para el diseño de los tests, la selección de los ítems, el método para afrontar el sesgo en los ítems, y el equating y notificación de los puntajes. El potencial de IRT para resolver este tipo de problemas es considerable. Sin embargo, para que puedan obtenerse los eventuales beneficios de un modelo de IRT resulta esencial que exista un ajuste entre el modelo y los datos del test que sean de interés. Un modelo de IRT que no se ajuste adecuadamente no generará parámetros invariantes.

En muchas aplicaciones de IRT de las que se informa en la bibliografía, el ajuste entre datos y modelo y las consecuencias de su

desajuste no han sido investigados adecuadamente. Como resultado de lo anterior se sabe menos de lo que podría suponerse, si se tiene en cuenta la voluminosa bibliografía existente sobre IRT, acerca de la conveniencia de determinados modelos de IRT para varias aplicaciones (Hambleton y otros, 1991, p. 53).

Hambleton y otros (1991) nos alertan aun más del peligro de confiarnos demasiado en los tests estadísticos, ya que éstos adolecen de un grave defecto: su sensibilidad al tamaño de la muestra de examinados. En lugar de ello los autores recomiendan que los juicios sobre el ajuste del modelo a los datos del test se basen en tres tipos de evidencia:

1. Validez de los supuestos del modelo para los datos del test.
2. Grado en que se obtienen las propiedades esperadas del modelo (por ejemplo, invarianza de los parámetros del ítem y de los parámetros de habilidad).
3. Exactitud de las predicciones del modelo empleando datos de tests reales y —si procede— simulados.

En las siguientes secciones de este trabajo se presentan los resultados de diversos tipos de análisis. El objetivo de estas investigaciones ha sido encontrar distintos tipos de evidencia en favor o en contra del ajuste de un modelo de IRT a los datos de la SweSAT.

1. LA APLICABILIDAD DE LOS MODELOS DE IRT A LOS SUBTESTS DE LA SWE SAT

IRT tiene un gran potencial para resolver muchos problemas en la aplicación de tests y en medición. Con todo, el éxito de determinadas aplicaciones de esta teoría no está asegurado por el solo hecho de procesar los datos del test por medio de uno de los programas computacionales (...). Los beneficios de los modelos de respuesta al ítem pueden obtenerse sólo cuando el ajuste entre el modelo y los datos de interés del test es satisfactorio (Hambleton y otros, 1991, p. 53).

Para investigar si un modelo de IRT podría ajustarse satisfactoriamente a cada uno de los 5 subtests DS, DTM, ERC, READ y WORD, se empleó una muestra aleatoria del 3% de los 85.506 examinados que rindieron la SweSAT en el trimestre de primavera de 1996. La muestra estaba compuesta de 2.461 alumnos: 1.349 mujeres y 1.112 hombres. Los resulta-

dos que obtuvieron estos examinados en cada subtest por separado proporcionaron los datos que se analizaron de distintas maneras.

La primera etapa consistió en realizar un análisis estándar de ítem por teoría clásica, cuyo resultado se presenta más abajo.

Análisis de ítem clásico

El análisis de ítem del subtest DS dentro del marco de la TCM arrojó un rango de valores p que iba de 0,40 a 0,81, y un rango de correlaciones biserials que iba de 0,25 a 0,70. El coeficiente de confiabilidad, alfa, fue $r = 0,82$.

El análisis de ítem del subtest DTM dentro del marco de la TCM arrojó un rango de valores p que iba de 0,28 a 0,82, y un rango de correlaciones biserials¹ que iba de 0,19 a 0,56. El coeficiente de confiabilidad, alfa, fue de $r = 0,72$.

El análisis de ítem del subtest ERC dentro del marco de la TCM arrojó un rango de valores p que iba de 0,28 a 0,82, y un rango de correlaciones biserials que iba de 0,19 a 0,56. El coeficiente de confiabilidad, alfa, fue de $r = 0,72$.

El análisis de ítem del subtest READ dentro del marco de la TCM arrojó un rango de valores p que iba de 0,34 a 0,84, y un rango de correlaciones biserials que iba de 0,21 a 0,45. El coeficiente de confiabilidad, alfa, fue de $r = 0,68$.

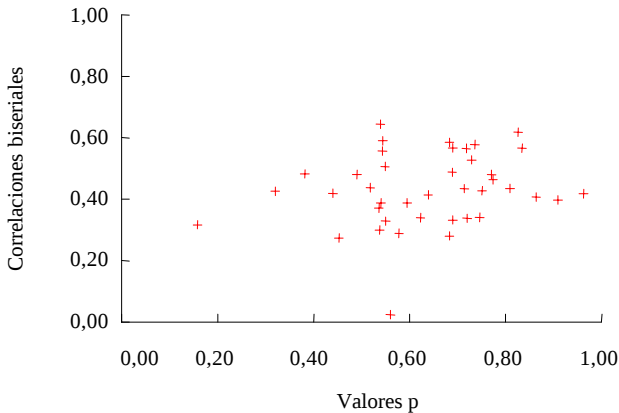
Por último, el análisis de ítem del subtest WORD dentro del marco de la TCM arrojó un rango de valores p que iba de 0,16 a 0,96, y un rango de correlaciones biserials que iba de 0,02² a 0,64. El coeficiente de confiabilidad, alfa, fue de $r = 0,85$.

Los rangos de correlaciones biserials indican que existe una considerable variación en el poder de discriminación de los ítems en la totalidad de los 5 subtests. A veces, sin embargo, el rango puede resultar engañoso debido a un par de “valores atípicos”. Además, altas correlaciones biserials se asocian a veces a ítems muy fáciles. Esos índices de discriminación no muestran realmente ítems eficaces, y por tanto los valores p fueron graficados contra las correlaciones biserials para todos los ítems en cada subtest. En la Figura 1 se muestra como ejemplo el gráfico para el subtest WORD.

¹ Las correlaciones biserials se calculan como la correlación entre el ítem y el puntaje total sin ese ítem.

² Había un ítem que se desviaba, el cual no operó adecuadamente; de ahí que la correlación biserial haya sido tan baja.

FIGURA 1: CORRELACIONES BISERIALES GRAFICADAS CONTRA VALORES P DE LOS 40 ÍTEMS DEL SUBTEST WORD



Los gráficos, que fueron similares para la totalidad de los 5 subtests, respaldaron el supuesto de que en realidad hubo variación en el poder discriminador de los ítems en todos los subtests. No pareció que existiera alguna conexión entre ítems fáciles y correlaciones biseriales altas. La conclusión fue que aparentemente se necesitaba un parámetro de discriminación de los ítems, y que por tanto un modelo IRT de un solo parámetro parecía inadecuado para los resultados de todos los subtests.

Con el fin de determinar si en los tests algunos alumnos habían adivinado las respuestas correctas, se estudió a los examinados que obtuvieron los resultados más bajos, escogiéndose a todos los que quedaron por debajo del primer percentil en cada subtest; los ítems difíciles se definieron como aquellos con valores p inferiores a 0,50.

Se estudiaron los resultados obtenidos por estos examinados de bajo desempeño en los ítems más difíciles de cada subtest, y se determinó que en el subtest DS los valores p para dichos alumnos en los 8 ítems más difíciles fueron:

p = 0,11; 0,30; 0,08; 0,14; 0,20; 0,13; 0,11; y 0,17

en el subtest DTM:

p = 0,26; 0,14; 0,11; 0,06; 0,17; 0,18; 0,11 y 0,20

en el subtest ERC:

p = 0,21; 0,24; 0,12; 0,22; 0,19; 0,15; y 0,35

en el subtest READ:

$p = 0,17; 0,16; 0,12; \text{ y } 0,15$

en el subtest WORD:

$p = 0,13; 0,01; 0,14; 0,01; 0,11; \text{ y } 0,22.$

Estos resultados indicaron que en cualquiera de los subtests difícilmente se puede excluir la posibilidad de que los alumnos hayan adivinado las respuestas, por lo que tampoco el modelo de dos parámetros parecía adecuado para ajustarse a los datos.

Análisis factorial

Un supuesto común a todos los modelos de IRT es que el grupo de ítems de los tests es unidimensional. Una medición bruta de la unidimensionalidad es el coeficiente de confiabilidad, alfa, ya que éste mide la coherencia interna de los ítems en un test. El coeficiente alfa varió entre 0,68 y 0,85 para los subtests. El coeficiente $r = 0,68$ indica que el subtest no es muy homogéneo, pero se trata sólo de una medición bruta. Un método más adecuado para evaluar la unidimensionalidad de un test es el análisis factorial (Hambleton y Rovinelli, 1986). Si el análisis factorial revela la existencia de sólo un factor dominante, sirve como respaldo para el argumento de la unidimensionalidad. Los resultados de los análisis factoriales fueron:

Para el subtest DS los análisis arrojaron tres factores con valores propios (*eigenvalues*) de 4,77; 1,21; y 1,09, respectivamente. La varianza explicada por el primer factor fue de 21,7%, y todos los ítems asignaban considerables cargas en el primer factor (entre 0,24 y 0,64).

Para el subtest DTM el resultado fue de 4 factores con valores propios de 3,3; 1,2; 1,1 y 1,0, respectivamente. La varianza explicada por el primer factor fue de 16,4%.

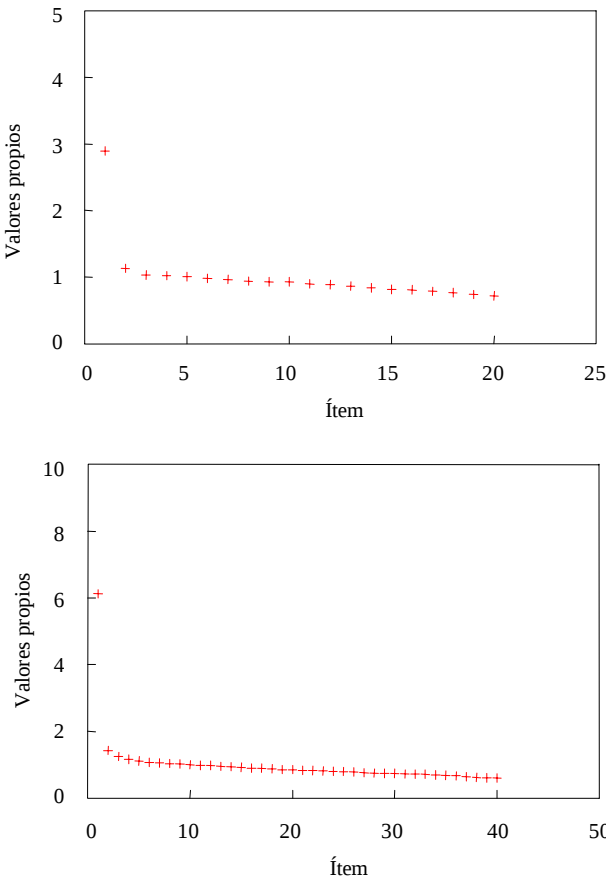
Para el subtest ERC el resultado fue también de 4 factores con valores propios de 3,8; 1,1; 1,0 y 1,0, respectivamente. La varianza explicada por el primer factor fue de 19,4 %.

Para el subtest READ el resultado fue de 5 factores con valores propios de 2,9; 1,1; 1,0; 1,0; y 1,0 respectivamente. La varianza explicada por el primer factor fue de 14,5%.

Para el subtest WORD, por último, el análisis factorial no rotado arrojó 9 factores con valores propios de: 6,1; 1,4; 1,2; 1,2; 1,1; 1,1; 1,0; 1,0 y 1,0 respectivamente. La varianza explicada por el primer factor fue de 15,4%.

Todos los valores propios fueron graficados, y los gráficos para los dos subtests con los primeros valores propios más pequeños, es decir, los subtests READ y WORD, aparecen en la Figura 2.

FIGURA 2: GRÁFICO DE VALORES PROPIOS PARA LOS SUBTESTS READ (ARRIBA) Y WORD (ABAJO)



En la Figura 2 se demuestra que, después de todo, parece haber un primer factor dominante en ambos subtests, ya que de acuerdo con Hambleton y Rovinelli (1986):

La cantidad de factores “significativos” se determina buscando el “codo” en el gráfico. La cantidad de valores propios a la izquierda del codo es normalmente interpretada como el número de factores significativos implícitos en el desempeño en un test (p. 289).

Aun cuando habría sido preferible que el valor de la varianza explicada por el primer factor fuera generalmente mayor, no resulta improbable o irracional suponer que existe un único factor en los datos de test para cualquiera de los subtests.

El modelo logístico de IRT con 3 parámetros

Se intentó ajustar los resultados de cada subtest al modelo logístico de IRT con 3 parámetros por medio del programa BILOGW (Mislevy y Bock, 1990). Cuando la cantidad de ítems es de 20 ó más, se incluyen como resultado del programa estadísticas aproximadas de chi-cuadrado para la calidad del ajuste de cada ítem. Para estos efectos los casos en la muestra de calibración se clasifican en intervalos sucesivos del continuo latente de acuerdo con las estimaciones de habilidad reescalada a media = 0 y desviación estándar = 1. Lo anterior permite obtener una prueba de ajuste razonable si la cantidad de ítems es lo suficientemente grande como para que una asignación de casos resulte precisa, y si el tamaño de la muestra es lo suficientemente grande como para contener 3 ó más intervalos. En estos estudios el menor número de ítems fue de 20, y la muestra de examinados fue extensa. La cantidad de intervalos utilizados fue de 10 para la mayoría de los ítems. Los siguientes fueron los resultados de los análisis de calidad del ajuste:

En cuanto al subtest DS, el resultado fue que para 11 ítems hubo un desajuste entre los datos y el modelo a nivel de $\alpha = 0,05$, y para 4 de esos ítems el desajuste fue considerable a nivel de $\alpha = 0,01$. El índice de confiabilidad fue de $r = 0,84$, que puede compararse con el coeficiente alfa en el análisis clásico, que fue de $r = 0,82$.

En cuanto al subtest DTM, el resultado fue que para 10 ítems hubo un considerable desajuste entre los datos y el modelo, y para 8 de esos ítems el desajuste fue significativo a nivel de $\alpha = 0,01$. El índice de confiabilidad fue de $r = 0,74$, en comparación con el coeficiente alfa, que fue de $r = 0,72$.

En cuanto al subtest ERC, el resultado fue de 14 ítems con un considerable desajuste entre los datos y el modelo, y para 7 de esos ítems el desajuste fue significativo a nivel de $\alpha = 0,01$. El índice de confiabilidad fue de $r = 0,80$, mientras que el coeficiente alfa fue de $r = 0,76$.

En cuanto al subtest READ, el resultado fue de 9 ítems con un considerable desajuste entre los datos y el modelo, y para 7 de esos ítems el desajuste fue significativo a nivel de $\alpha = 0,01$. El índice de confiabilidad reportado fue de $r = 0,72$, y el coeficiente alfa fue de $r = 0,68$.

Y finalmente el resultado de los análisis de calidad de ajuste correspondiente al subtest WORD fue que para 8 de los ítems hubo un considerable desajuste entre los datos y el modelo a nivel de $\alpha = 0,05$, y para uno de estos 8 ítems el desajuste fue significativo a nivel de $\alpha = 0,01$. El índice de confiabilidad registrado fue de $r = 0,87$, mientras que el coeficiente de confiabilidad, alfa, fue de $r = 0,85$.

Análisis residuales

Se realizaron otros tipos de análisis de calidad del ajuste mediante el programa RESID (Rogers, 1994). Al efectuarse estos análisis los examinados se clasifican primero en categorías de habilidad. La cantidad de niveles de habilidad fue especificada en 8, y se calcularon las proporciones observadas de examinados en cada categoría de habilidad que respondieron el ítem correctamente. Las proporciones esperadas de respuestas correctas para cada intervalo de habilidad se obtuvieron calculando la probabilidad de éxito en el ítem en cada nivel de habilidad. Luego se determinaron los valores residuos (observados menos esperados) y los residuos estandarizados. El programa contiene asimismo estadísticas de ajuste chi-cuadrado para cada ítem como resultado.

Los siguientes fueron los resultados de los análisis RESID:

En cuanto al subtest DS, las diferencias entre valores observados y esperados fueron estadísticamente significativas para 2 ítems, ambos en el nivel 0,05.

En cuanto al subtest DTM, las diferencias fueron estadísticamente significativas para 6 ítems, 5 de ellos en el nivel 0,01.

En cuanto al subtest ERC, las diferencias fueron estadísticamente significativas para 7 ítems, 2 de ellos en el nivel 0,01.

En cuanto al subtest READ, las diferencias fueron estadísticamente significativas para 5 ítems, 2 de ellos en el nivel 0,01.

Por último, en cuanto al subtest WORD, las diferencias fueron estadísticamente significativas para 6 ítems, uno de ellos en el nivel 0,01.

Los residuos permiten comparar el desempeño pronosticado y el real. Los residuos brutos corresponden a las diferencias entre desempeño esperado y observado en un ítem en un nivel de desempeño especificado. Los residuos estandarizados (RE) tienen en cuenta el error muestral asociado a cada nivel de desempeño, al igual que la cantidad de examinados en ese nivel de desempeño en particular. Cuando el modelo se ajusta a los datos podría esperarse que los RE fueran pequeños y se distribuyeran alea-

toriamente alrededor de 0. Dentro del marco de la teoría de la regresión resulta común suponer que la distribución de los RE es aproximadamente normal. En la Tabla 1 se entrega un resumen de los RE de los análisis de calidad de los ajustes.

TABLA 1: VALORES ABSOLUTOS DE LOS RESIDUOS ESTANDARIZADOS (%) PARA LOS 5 SUBTESTS

RE	DS	DTM	ERC	READ	WORD
0-1	72,16	65,63	62,50	68,75	70,31
1-2	22,16	25,00	31,25	24,38	26,25
2-3	5,11	7,50	5,00	6,88	3,13
>3	0,57	1,88	1,25	0,00	0,31

Los resultados que aparecen en la Tabla 1 demuestran que aun cuando las distribuciones para los subtests DTM y ERC son hasta cierto punto demasiado uniformes, todas las distribuciones de RE se aproximan bastante a la distribución normal que se supone proporciona un sólido respaldo al ajuste del modelo a los datos.

Hambleton y otros (1991) han formulado las siguientes recomendaciones en cuanto a la evaluación de ajuste del modelo a los datos:

Al evaluar el ajuste del modelo a los datos, el mejor enfoque implica: a) diseñar y realizar una diversidad de análisis destinados a detectar los tipos de desajuste esperados; b) considerar cuidadosamente la serie completa de resultados; y c) emitir un juicio acerca de la conveniencia del modelo para la aplicación proyectada. Los análisis deberían incluir investigaciones de supuestos del modelo, del grado en que se obtienen las características deseadas del modelo, y de las diferencias entre predicciones del modelo y datos reales. Se pueden realizar tests estadísticos, pero es preciso tener cuidado al interpretar la información estadística. El número de investigaciones que pueden llevarse a cabo es casi ilimitado (p. 74).

Para el subtest DS se efectuó una comparación entre los parámetros de los ítems, estimados en dos muestras distintas de alumnos que rindieron el examen. La correlación entre valores *b* fue de $r = 0,95$, y la correlación entre valores *a* fue de $r = 0,72$.

Para el subtest DTM se realizó una comparación entre parámetros estimados dentro de la IRT e índices calculados al interior de la TCM. La

correlación entre los valores *b* estimados y los valores *p* calculados fue de $r = -0,94$, y la correlación entre valores *a* estimados y correlaciones biseriales calculadas (r_{bis}) fue de $r = 0,82$. La correlación entre puntajes del test y parámetros estimados de habilidad fue de $r = 0,95$. Se efectuó otra comparación entre parámetros estimados en alumnos y alumnas que rindieron el test, y el resultado fue que la correlación entre valores *b* estimados en hombres y mujeres examinados fue de $r = 0,89$, y entre valores *a* de $r = 0,91$.

Se efectuaron comparaciones análogas para el subtest ERC. La correlación entre valores *b* estimados y valores *p* calculados en este subtest fue de $r = -0,88$, y entre valores *a* estimados y r_{bis} calculada fue de $r = 0,73$. La correlación entre valores *b* estimados en hombres examinados y valores *b* estimados en mujeres examinadas fue de $r = 0,93$, y la correlación entre valores *a* fue de $r = 0,83$.

Las mismas comparaciones se efectuaron para los subtests READ y WORD. Para el subtest READ la correlación entre valores *p* y *b* fue de $r = -0,98$, y para el subtest WORD dicha correlación fue de $r = -0,74$. Para el subtest READ la correlación entre r_{bis} y valores *a* fue de $r = 0,84$, mientras que para el subtest WORD fue de $r = 0,82$. La correlación entre valores *b* estimados en hombres y valores *b* calculados en mujeres en el subtest READ fue de $r = 0,92$, y en el subtest WORD esta correlación fue de $r = 0,85$. La correlación entre valores *a* estimados en hombres examinados y los valores *a* calculados en mujeres examinadas fue de $r = 0,76$ en el subtest READ y de $r = 0,77$ en el subtest WORD. Los resultados de estas comparaciones entre grupos se presentan en la Tabla 2.

TABLA 2: RELACIONES ENTRE PARÁMETROS DE LOS ÍTEMS ESTIMADOS DENTRO DE LA IRT E ÍNDICES DE ÍTEMS CALCULADOS DENTRO DE LA TCM, Y ENTRE PARÁMETROS DE LOS ÍTEMS DE IRT ESTIMADOS EN HOMBRES Y MUJERES EXAMINADOS

Correlación	DTM	ERC	READ	WORD
valores <i>p</i> y <i>b</i>	-0,94	-0,88	-0,98	-0,74
r_{bis} y valores <i>a</i>	0,82	0,73	0,84	0,82
M y H est. <i>b</i>	0,89	0,93	0,92	0,85
M y H est. <i>a</i>	0,91	0,83	0,76	0,77

M: mujeres; H: hombres.

1.1. Análisis

Los resultados de estos intentos por ajustar separadamente un modelo logístico de IRT de 3 parámetros a cada uno de los 5 subtests de la SweSAT son algo variados. Los resultados de los análisis iniciales con la TCM respaldaron la necesidad de aplicar un modelo de 3 parámetros en la totalidad de los 5 subtests. Los análisis factoriales respaldaron la tesis de la unidimensionalidad en el subtest DS, para el cual el primer factor podía explicar el 21,7% de la varianza del test. En cuanto a los otros subtests el respaldo de los análisis factoriales para la unidimensionalidad fue más débil: 19,4% de la varianza explicada por el primer factor para el subtest ERC; 16,4% para el subtest DTM; 15,4% para el subtest WORD; y 14,5% para el subtest READ. Cuando la varianza explicada es inferior al 20% no se tiene la certeza para poder afirmar que el test es unidimensional.

El supuesto de independencia local también resulta problemático al menos para 3 de los subtests del SweSAT. El subtest READ consta de 4 textos con 5 preguntas en relación con cada texto. Pese a que estos 5 ítems son independientes entre sí, pertenecen al mismo texto. El subtest DTM está compuesto por 10 figuras, tablas y mapas con dos preguntas para cada representación gráfica. El subtest ERC está formado por una cantidad variable de textos con 2 a 5 preguntas para cada uno. Existen dudas en cuanto a si los ítems son en verdad localmente independientes en los tests de este formato.

Otro problema se observa en los distintos resultados obtenidos en las 2 pruebas estadísticas. Se descubrió que por lo general la cantidad de ítems que presentaban un desajuste estadísticamente significativo entre los datos y el modelo era mayor aplicando el programa BILOG que aplicando el programa RESID, y ello podría explicarse, al menos en parte, por el hecho de que el BILOG divide en más grupos de habilidad que el RESID. Sin embargo, se descubrió que mientras algunos de los ítems exhibían un considerable desajuste entre los datos y el modelo empleando el programa RESID, ello no ocurría al aplicar el BILOG. Este fenómeno se verificó en 2 ítems del subtest DS, 2 del subtest DTM, uno del subtest ERC, 2 del subtest READ, y 3 del subtest WORD.

Hasta ahora los análisis efectuados no han confirmado totalmente ni rechazado de plano el ajuste entre el modelo logístico de IRT de 3 parámetros y los datos del SweSAT.

2. COMPARACIÓN ENTRE LOS ANÁLISIS DE ÍTEMS BASADOS EN LA TEORÍA DE RESPUESTA AL ÍTEM Y EN LA TEORÍA CLÁSICA

Tal como ocurre en todas las evaluaciones altamente decisivas, la serie de pruebas piloto o pretests de los ítems del SweSAT constituyen una etapa crucial del proceso de elaboración de este test. El pretest tiene diversos objetivos (Henrysson, 1971), de los cuales los más importantes para el SweSAT son:

- Determinar la dificultad de cada ítem de modo que se pueda hacer una selección, la cual indicará el nivel de dificultad del subtest, que es similar a anteriores versiones de éste.
- Identificar ítems débiles o defectuosos con distractores que no funcionan.
- Determinar el poder de cada ítem para discriminar entre examinados de buen y mal desempeño en la variable de logro medida.
- Identificar ítems sesgados (por género).

Desde que la SweSAT se aplicó por primera vez en 1977, la elaboración y la compilación del test, lo mismo que el equating de las diferentes formas de éste entre una administración y la próxima se ha basado en la TCM. Sobre la base de los datos obtenidos en los pretests, los ítems son rechazados o seleccionados para el test definitivo, y las estadísticas que se usan en el análisis de ítems son:

Valores p de los ítems.

Valores p de los distractores.

Correlaciones biserials (r_{bis}).

Valores p de hombres y mujeres.

La regresión del test por ítem.

La principal limitación de la TCM a este respecto es que la estadística de la persona (es decir el puntaje obtenido en el test) depende de la muestra de ítems (es decir, del test), y las estadísticas de los ítems dependen de la muestra de examinados. El argumento esencial para preferir el uso de modelos de IRT al empleo de procedimientos de TCM es que la IRT debería generar como resultado medidas independientes de las muestras. Con IRT una persona debería teóricamente recibir la misma estimación de habilidad, independientemente del test rendido, y las estadísticas de los ítems deberían mantenerse estables en distintos grupos de individuos. De allí que la gran

ventaja de IRT sea la invarianza de los parámetros de los ítems. Una desventaja de la IRT es que para estimar los parámetros los tamaños de las muestras tienen que ser grandes.

IRT ha sido objeto de intensas investigaciones por parte de los psicometristas, y además se han publicado numerosos libros y artículos (Fan, 1988). No obstante, los estudios empíricos disponibles se han centrado principalmente en diversas aplicaciones de IRT, y muy pocos han comparado efectivamente la TCM e IRT en cuanto al análisis de ítems y el diseño de tests. Continúa Fan (1998):

Resulta hasta cierto punto sorprendente que los estudios empíricos en que se examinan y/o comparan las características de invarianza de estadísticas de ítems dentro de ambos marcos de medición sean tan escasos. Tal parece que en la comunidad de medición se ha dado por sentada la superioridad de IRT sobre TCM a este respecto, y no se ha estimado necesaria ninguna indagación empírica. El silencio empírico en torno a este tema parece ser anómalo (p. 361).

Desde el segundo trimestre de 1996 el pretest de los ítems para la SweSAT se ha realizado en conexión con la administración regular de la prueba, lo cual significa que la muestra de examinados a la cual se aplica el pretest ha sido extraída del verdadero universo de examinados y está compuesta por 1.000 individuos como mínimo. Este nuevo procedimiento para realizar el pretest permitiría emplear la IRT para analizar ítems y compilar nuevas versiones de la prueba.

2.1. Objetivo

La finalidad del presente estudio fue comparar las estadísticas de los ítems dentro del marco de la TCM con los parámetros de los ítems dentro del marco de IRT, y analizar la estabilidad de las dos series de características de los ítems desde la etapa de pretest hasta la del test regular. Específicamente se formularon las siguientes preguntas:

1. ¿De qué manera se comparan los índices de dificultad de los ítems determinados dentro del marco de la TCM con los parámetros de dificultad de los ítems calculados por IRT?:

- a) ¿para los datos del pretest?,
- b) ¿para los datos del test regular?

2. ¿De qué manera se comparan los índices de discriminación de los ítems determinados dentro del marco de la TCM con los parámetros de discriminación de los ítems estimados por IRT?:

- a) ¿para los datos de pretests?,
 - b) ¿para los datos del test regular?
3. ¿Cuál es el grado de estabilidad de los índices de ítems dentro de la TCM desde los datos de pretest hasta los datos del test regular?
4. ¿Cuál es el grado de estabilidad de los parámetros de ítems dentro de IRT desde los datos del pretest hasta los datos del test regular?

2.2. Método

Teoría clásica de la medición

En el test regular aplicado durante el segundo trimestre de 1997 se incluían 20 ítems de WORD, 16 de READ y 14 de ERC, los que ya habían sido pretesteados durante el segundo trimestre de 1996. Para estos ítems se calcularon los valores p y las correlaciones biseriales. Se estimaron los mismos índices para los correspondientes ítems en los datos del pretest, y se compararon los valores.

Teoría de respuesta al ítem

Las 5 combinaciones de pretests para el WORD, en las cuales habían sido diseminados los ya mencionados 20 ítems de WORD, se ejecutaron con el programa BILOGW junto con el test regular de WORD desde el segundo trimestre de 1996, y se estimaron los parámetros a , b y c . A partir del segundo trimestre de 1997 el subtest WORD también se ejecutó en BILOGW y se calcularon los parámetros de los ítems. Por último, se compararon los parámetros estimados para los 20 ítems comunes.

Las 8 versiones de pretests para el READ se ejecutaron en BILOGW a contar del segundo trimestre de 1996 junto con el subtest regular de READ, y se estimaron los parámetros a , b y c . El subtest de READ se ejecutó en BILOGW desde el segundo trimestre de 1997, y se estimaron los parámetros de los ítems. Se registraron y compararon las estimaciones de los parámetros para los 16 ítems comunes.

Las 4 versiones de pretests de ERC fueron ejecutadas en BILOGW a partir del segundo trimestre de 1996 junto con el subtest regular de ERC, y se calcularon los parámetros a , b y c . A partir del segundo trimestre de 1997 el subtest de ERC se ejecutó en BILOGW y se estimaron los parámetros de los ítems. Se registraron y compararon las estimaciones de los parámetros de los ítems para los 14 ítems comunes.

2.3. Resultados

Un problema que se presenta al analizar la estabilidad de los parámetros de los ítems es que el pretest tiene dos objetivos. Uno es el de obtener información sobre el nivel de dificultad y el poder de discriminación de los ítems para así estar en condiciones de compilar tests de dificultad análoga. El otro objetivo es asegurarse de que todos los ítems funcionen de manera satisfactoria. Si un ítem no está funcionando con la suficiente eficacia, será modificado o excluido. Si se introducen cambios importantes, el ítem será pretesteado una vez más antes de ser puesto en práctica, pero si los cambios son menores será aplicado en el test regular. Tales cambios, sin embargo, dan a entender que los ítems no son exactamente los mismos en la versión del pretest que en la del test regular. Otro problema es que los ítems pueden ser dispuestos en distinto orden en el cuadernillo del pretest y en el del test regular, y que los ítems tienden a tornarse más difíciles al final del cuadernillo. En las Tablas 3 a 5 se entrega la ubicación del ítem en el cuadernillo del pretest al igual que en el del test regular, y los ítems en los cuales se introdujeron cambios menores entre ambas etapas están marcados con un *.

El subtest WORD

En la Tabla 3 se muestran los índices de dificultad y discriminación calculados dentro del marco de la TCM para 20 ítems de WORD en el pretest, al igual que en test regular. En la misma tabla se presentan los parámetros de dificultad y discriminación estimados dentro del marco de IRT.

La correlación entre valores p calculada en datos de pretests y los valores p calculados en datos del test regular fue de $r = 0,93$.

La correlación entre valores b estimados en datos de pretests y valores b estimados en datos del test regular fue de $r = 0,92$.

La correlación entre valores p y b fue de $r = -0,93$, tanto para los datos de pretests como para los del test regular.

La correlación entre r_{bis} calculadas en datos de pretests y r_{bis} calculadas en datos del test regular fue de $r = 0,81$.

La correlación entre valores a estimados en datos de pretests, y de valores a estimados en datos del test regular fue de $r = 0,74$.

La correlación entre la r_{bis} de discriminación de los ítems y el parámetro de discriminación de los ítems a fue de $r = 0,65$ para los datos de pretests y de $r = 0,64$ para los datos del test regular.

TABLA 3: CARACTERÍSTICAS DE LOS ÍTEMS PARA 20 ÍTEMS DE WORD CALCULADAS DENTRO DEL MARCO DE LA TCM Y ESTIMADAS DENTRO DEL MARCO DE IRT

Ítem N°		Dificultad TCM		Dificultad IRT		Discrimin. TCM		Discrimin. IRT	
pretest	test reg.	pretest	test reg.	pretest	test reg.	pretest	test reg.	pretest	test reg.
8	1*	0,73	0,71	-0,43	-0,43	0,60	0,58	1,29	1,14
20	4*	0,79	0,72	-0,96	-0,64	0,46	0,41	0,73	0,62
39	5*	0,78	0,74	-1,94	-1,19	0,25	0,33	0,32	0,43
18	9	0,68	0,71	-0,25	-0,59	0,44	0,43	0,71	0,63
36	10*	0,75	0,72	-0,92	-0,81	0,50	0,53	0,72	0,78
27	11*	0,80	0,82	-1,49	-1,79	0,35	0,35	0,48	0,46
36	15*	0,71	0,58	-0,55	0,11	0,40	0,37	0,59	0,55
14	16*	0,65	0,70	-0,02	-0,65	0,44	0,48	0,81	0,72
5	19	0,46	0,42	0,46	0,71	0,42	0,37	0,55	0,50
16	23	0,65	0,62	0,11	0,08	0,35	0,40	0,58	0,72
38	24	0,58	0,56	0,08	0,16	0,47	0,30	0,75	0,40
12	25	0,51	0,59	0,17	-0,02	0,58	0,58	0,97	1,13
24	27	0,69	0,66	0,37	0,33	0,36	0,35	1,07	0,83
4	28*	0,53	0,44	0,35	0,51	0,56	0,52	1,55	1,04
4	29	0,42	0,42	1,16	1,08	0,33	0,26	0,71	0,61
5	35*	0,31	0,38	1,59	0,89	0,32	0,46	0,45	0,95
37	36*	0,71	0,62	-0,37	-0,07	0,43	0,44	0,70	0,70
6	38*	0,41	0,46	1,25	0,49	0,31	0,37	0,70	0,48
28	39*	0,27	0,40	1,61	1,22	0,28	0,32	1,13	1,18
39	40*	0,31	0,31	2,27	2,45	0,23	0,21	0,42	0,45

El subtest READ

La correlación entre valores p de los ítems en las versiones de pretests, y valores p de los ítems correspondientes en el test regular fue de $r = 0,78$.

La correlación entre valores b calculada en datos de pretests y valores b estimados en datos del test regular fue de $r = 0,55$.

La correlación entre valores p calculados en datos del pretest y valores b estimados en datos de pretests fue de $r = -0,90$.

La correlación entre TCM e IRT en cuanto a los índices de dificultad en datos del test regular fue de $r = -0,92$.

La correlación entre la r_{bis} de los ítems en las versiones de pretests y la r_{bis} de los ítems correspondientes en el test regular fue de $r = 0,66$.

La correlación entre valores a estimados en datos de pretests y valores a estimados en datos del test regular fue de $r = 0,54$.

La correlación entre los índices de discriminación de TCM e IRT en los datos de pretests fue de $r = 0,35$. La correlación entre índices de discriminación de CTT e IRT en datos del test regular fue de $r = 0,78$.

TABLA 4: CARACTERÍSTICAS DE 16 ÍTEMS DE READ CALCULADAS DENTRO DEL MARCO DE LA TCM Y ESTIMADAS DENTRO DEL MARCO DE IRT

Ítem N°		Dificultad TCM		Dificultad IRT		Discrimin. TCM		Discrimin. IRT	
pretest	test reg.	pretest	test reg.	pretest	test reg.	pretest	test reg.	pretest	test reg.
5	5	0,74	0,74	-0,61	-0,69	0,30	0,32	0,50	0,59
7	6*	0,30	0,68	2,16	-0,52	0,23	0,36	0,46	0,64
6	7	0,78	0,71	-1,16	-0,45	0,37	0,38	0,54	0,73
8	8	0,80	0,81	-1,31	-1,09	0,40	0,35	0,59	0,63
17	9	0,64	0,81	0,37	-0,88	0,37	0,43	0,85	0,90
20	10	0,52	0,69	1,32	0,27	0,25	0,22	0,64	0,42
17	11	0,61	0,75	1,42	-1,20	0,18	0,11	0,52	0,37
20	12	0,45	0,52	1,92	0,71	0,17	0,33	0,72	0,86
14	13	0,59	0,66	0,48	-0,14	0,35	0,36	0,83	0,69
15	14*	0,36	0,50	1,85	0,74	0,24	0,30	0,53	0,72
14	15*	0,24	0,30	1,67	1,08	0,34	0,43	0,87	1,23
16	16*	0,28	0,57	2,54	0,33	0,17	0,28	0,59	0,54
17	17*	0,35	0,53	1,33	0,51	0,35	0,36	0,72	0,69
19	18	0,63	0,76	0,82	-0,77	0,29	0,32	0,76	0,60
19	19	0,56	0,65	0,84	0,04	0,29	0,31	0,69	0,70
20	20	0,57	0,60	0,38	0,22	0,34	0,34	0,57	0,91

El subtest ERC

TABLA 5: CARACTERÍSTICAS DE 14 ÍTEMS DE ERC CALCULADAS DENTRO DEL MARCO DE LA TCM Y ESTIMADAS DENTRO DEL MARCO DE IRT

Ítem N°		Dificultad TCM		Dificultad IRT		Discrimin. TCM		Discrimin. IRT	
pretest	test reg.	pretest	test reg.	pretest	test reg.	pretest	test reg.	pretest	test reg.
1	1	0,33	0,38	1,65	1,31	0,30	0,33	0,59	0,79
2	2	0,72	0,66	0,20	0,34	0,34	0,33	0,83	0,76
3	3	0,35	0,29	1,70	1,74	0,28	0,29	-0,72	0,71
4	4	0,41	0,47	0,83	0,54	0,45	0,47	0,76	0,83
5	5*	0,62	0,73	0,78	0,35	0,16	0,54	0,27	1,05
1	6	0,77	0,78	-0,90	-0,62	0,62	0,54	1,00	1,10
2	7	0,54	0,50	0,09	0,49	0,48	0,46	0,67	0,81
3	8	0,53	0,53	0,77	0,63	0,37	0,42	0,90	0,99
11	9*	0,41	0,60	0,81	-0,07	0,41	0,38	0,56	0,55
5	10	0,67	0,58	-0,52	0,19	0,57	0,51	0,84	1,02
14	12	0,60	0,51	-0,13	-0,10	0,51	0,46	0,71	0,73
13	13	0,68	0,62	-0,27	-0,03	0,56	0,56	0,96	1,10
14	14	0,65	0,65	-0,40	-0,19	0,57	0,52	0,85	0,91
10	15	0,77	0,74	-0,85	-0,52	0,52	0,52	0,80	0,95

El promedio de los valores p fue de 0,58 para los datos de pretests, al igual que para los del test regular. La correlación entre valores p de los ítems en las versiones de pretests y de valores p de los correspondientes ítems en el test regular fue de $r = 0,86$.

El promedio de los valores b fue de 0,27 para los ítems del pretest y de 0,29 para los ítems del test regular. La correlación entre valores b estimados en datos del pretest y valores b estimados en datos del test regular fue de $r = 0,88$.

La correlación entre la r_{bis} de los ítems en las versiones del pretest y la r_{bis} de los ítems correspondientes en el test regular fue de $r = 0,57$.

La correlación entre valores a estimados en datos del pretest y valores a estimados en datos del test regular fue de $r = 0,34$.

Para los 12 ítems, que no habían sido modificados entre los pretests y el test regular, la correlación fue de $r = 0,95$ para los valores p , y $r = 0,96$ y $\rho = 0,94$ para r_{bis} .

Para los 12 ítems invariables la correlación entre valores b del pretest y del test regular fue de $r = 0,96$, y entre valores a , de $r = 0,82$.

La correlación entre dificultades de TCM y dificultadores de IRT fue de $r = -0,90$ para los datos del pretest, al igual que para los datos del test regular. La correlación entre discriminaciones calculada dentro del marco de la TCM y estimada dentro del marco de la IRT fue de $r = 0,74$ para los datos del pretest y de $r = 0,76$ para los datos del test regular.

Los resultados para los 3 subtests se resumen en la Tabla 6.

TABLA 6: ESTABILIDAD DE ÍNDICES DE ÍTEMS DE TCM Y DE PARÁMETROS DE ÍTEMS DE IRT DESDE VERSIONES DEL PRETEST HASTA LA VERSIÓN DEL TEST REGULAR. RELACIONES ENTRE ÍNDICES DE ÍTEMS DE TCM Y PARÁMETROS DE ÍNDICES DE IRT EN EL PRETEST Y EN EL TEST REGULAR

subtest	pretests y test reg.		pretests y test reg.		valores p y b		r_{bis} y valores a	
	p	b	r_{bis}	a	pre	reg.	pre	reg.
ERC	0,86 (0,95)	0,88 (0,92)	0,57	0,74	-0,90	-0,90	0,74	0,76
READ	0,78	0,55	0,66	0,54	-0,90	-0,92	0,35	0,68
WORD	0,93	0,92	0,81	0,74	-0,93	-0,93	0,64	0,65

2.4. Análisis

Para los subtests WORD y READ (el modelo logístico de 3 parámetros), en ninguno de los ítems de pretests se detectó un desajuste considerable con respecto al modelo logístico de 3 parámetros. Sin embargo, para 3 ítems del pretest de ERC se observó un desajuste, el cual era significativo a nivel de $\alpha = 0,01$. En los subtests regulares hubo un ítem en el subtest WORD, uno en el subtest READ, y 2 en el subtest ERC, en los cuales se observó un desajuste entre modelo y los datos, que era considerable a nivel de $\alpha = 0,01$. Estos ítems eran el N° 10 en el subtest regular de WORD, el N° 11 en el subtest regular de READ, y los N°s 9 y 14 en el subtest ERC.

La conclusión general de los estudios es que la concordancia entre los resultados de los análisis de ítems dentro de los dos marcos distintos, TCM e IRT, fue razonablemente aceptable. La correlación entre dificultades de los ítems para las versiones del test regular fue de $r = -0,93$ para el subtest WORD, $r = -0,92$ para el subtest READ, y $r = -0,90$ para el subtest ERC.

En cuanto a la estabilidad de los datos desde la etapa de pretests hasta la del test regular no hubo grandes diferencias entre ambas teorías. Para el subtest WORD la concordancia entre dificultades en los pretests y en el test regular fue de $r = 0,93$ dentro de la TCM, y de $r = 0,92$ dentro de IRT. Para el subtest READ la correlación entre dificultades de TCM fue de $r = 0,78$, y entre dificultades de IRT la correlación fue de $r = 0,55$. Para el subtest ERC la correlación al interior de TCM fue de $r = 0,87$ y al interior de IRT de $r = 0,88$. En términos generales, las correlaciones entre dificultades de ítems en los datos de pretests y del test regular fueron en realidad más altas para los índices de TCM que para los parámetros de IRT.

Puesto que en teoría la IRT difiere considerablemente de la TCM, y posee algunas ventajas competitivas cruciales con respecto a esta última, parece razonable esperar que existan apreciables diferencias entre las estadísticas sobre personas e ítems basadas en IRT y en la TCM. En teoría esas relaciones no son del todo claras, excepto que los dos tipos de estadísticas deberían estar monótonamente relacionadas bajo ciertas condiciones (Crocker y Algina, 1986; Lord, 1980), pero dichas relaciones rara vez han sido objeto de investigaciones empíricas, y por ende en gran parte son desconocidas (Fan, 1998, p. 360).

La conclusión general derivada de estas comparaciones es que la predicción desde los datos de pretests hasta los datos del test regular es aceptable, pero que ello se puede aplicar tanto a la TCM como a IRT.

A decir verdad, las predicciones formuladas dentro del marco de la TCM fueron por lo general más correctas que las planteadas dentro del marco de IRT. Los parámetros de ítems de IRT no fueron completamente invariantes. Puesto que los grupos en los cuales se aplicaron los pretests constituían muestras amplias y representativas del universo real de examinados, era dable esperar este resultado. Con todo, el dilema es que para lograr estimar los parámetros de ítems dentro del marco de IRT resulta imperioso contar con muestras amplias, y cuando éstas son lo suficientemente extensas, los índices de ítems dentro del marco de la TCM son asimismo muy estables.

Lo que suele considerarse la principal desventaja de la TCM es que las estadísticas de ítems, tales como la dificultad y la discriminación, dependen de la muestra particular de examinados de la cual se obtienen. Con frecuencia se estima que la principal ventaja teórica de IRT radica en la invarianza de las correspondientes estadísticas de los ítems. La invarianza de los parámetros de ítems en todos los grupos es una de las características más importantes de la teoría de respuesta al ítem (Lord, 1980, p. 35). En los estudios citados en este trabajo los parámetros de los ítems estimados dentro del marco de la IRT no fueron superiores, en cuanto a la invarianza entre grupos, a las estadísticas derivadas dentro del marco de la TCM. El problema puede ser que para lograr esta invarianza de parámetros de IRT, el ajuste entre el modelo y los datos debe ser perfecto. Desgraciadamente no existen criterios objetivos sobre el ajuste del modelo a los datos, pero de acuerdo con Hambleton y otros (1991) “...invarianza y ajuste del modelo a los datos son conceptos equivalentes” (p. 24).

3. APLICABILIDAD DE LA IRT A LA SWESAT: EL TEST TOTAL

La SweSAT es puntuada de acuerdo con la teoría clásica de la medición (TCM); el puntaje bruto para cada examinado corresponde al número de ítems respondidos correctamente. Todos los ítems tienen el formato de selección múltiple y son puntuados de manera dicotómica, es decir “1” para una respuesta correcta y “0” para una incorrecta. Se ha descubierto que en este tipo de ítems los examinados difieren en cuanto a su tendencia a adivinar respuestas, o a omitir ítems cuya respuesta correcta ignoran, lo cual puede generar una varianza irrelevante. Si bien se han inventado varios métodos para rectificar el efecto de adivinar las respuestas, los estudios empíricos no han respaldado el empleo de métodos de corrección (Crocker y Algina, 1986, p. 403). A quienes rinden la SweSAT se los insta a marcar una respuesta en todos los ítems. El puntaje total del test se ecualiza y se

transforma en un puntaje normado, el cual se utiliza en el proceso de selección para ingreso a la educación superior.

La SweSAT consta de 5 subtests que miden el conocimiento de vocabulario (WORD), el pensamiento lógico (DS), la comprensión de lectura en sueco (READ), la comprensión de lectura en inglés (ERC), y la habilidad para interpretar diagramas, tablas y mapas (DTM). En estudios anteriores se investigó la aplicabilidad de un modelo logístico de IRT de 3 parámetros a cada subtest (Stage, 1996, 1997a, 1997b, 1997c, 1997d). El resultado de esos estudios no fue una confirmación ni una refutación del ajuste entre el modelo logístico de IRT de 3 parámetros y los datos de la SweSAT.

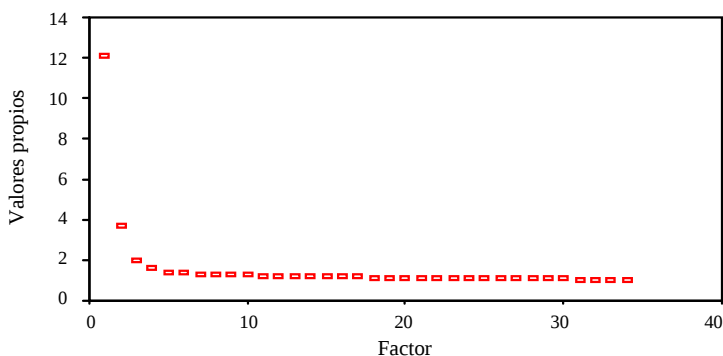
Puesto que es el puntaje total del test lo que indica el resultado obtenido por quienes rindieron la SweSAT, la aplicabilidad de un modelo de IRT adecuado al test total resulta importante. En el presente estudio el objetivo era investigar si un modelo logístico de IRT de 3 parámetros podía aplicarse exitosamente al test total.

Se escogió el modelo de 3 parámetros porque había resultado ser el más adecuado para cada uno de los subtests. No había motivos para creer que el test total debiera diferenciarse de sus partes en cuanto a la necesidad de un parámetro de discriminación y de un parámetro de pseudoadivinanza.

Unidimensionalidad

Un análisis factorial no rotado de los puntajes de los tests totales arrojó un primer factor de 12,1, un segundo factor de 3,7, un tercer factor de 2,0, y 34 factores por sobre 1,0. El primer factor explicaba el 9,9% de la varianza, el segundo factor el 3,0%, y los restantes factores entre el 1,7% y el 0,8% cada uno. En la figura 3 se presenta un gráfico de los valores propios.

FIGURA 3: GRÁFICO DE VALORES PROPIOS PARA LA TOTALIDAD DE LA SWESAT



En la Figura 3 puede apreciarse que aun cuando existe una especie de “codo” en el gráfico, y el primer factor es dominante, el supuesto de unidimensionalidad resulta incierto, ya que el segundo factor es hasta cierto punto demasiado fuerte.

También se efectuó un análisis factorial a nivel de subtests, el cual arrojó un primer factor con valor propio de 3,05; un segundo factor con valor propio de 0,83; un tercero con 0,41; un cuarto con 0,38, y un quinto con 0,35. El primer factor explicaba el 61% de la varianza. Este análisis proporcionó mayor respaldo al supuesto de la unidimensionalidad.

El modelo logístico de 3 parámetros

El programa BILOGW se empleó para ajustar los resultados de la SweSAT al modelo logístico de IRT de 3 parámetros. El resultado del análisis de calidad del ajuste de los ítems fue que para 67 de ellos había un desajuste entre datos y modelo que era significativo a nivel de $\alpha = 0,05$, y para 44 de ellos el desajuste era considerable a nivel de $\alpha = 0,01$. El programa se ejecutó en 28.505 resultados del test, y puesto que las pruebas estadísticas de ajuste del modelo son muy sensibles al tamaño muestral, tales resultados eran esperables. Esto es lo que Hays (1969) llama la falacia de evaluar un resultado teniendo en cuenta sólo la significación estadística:

Es posible lograr que prácticamente cualquier estudio arroje resultados significativos si empleamos suficientes sujetos, sin importar cuán absurdo pueda ser el contenido (p. 326).

El programa también fue ejecutado en dos muestras aleatorias distintas de 1.000 resultados de tests. Para la primera muestra la cantidad de ítems en los cuales se observó un significativo desajuste había disminuido a 7, y en 4 de dichos ítems el desajuste era considerable a nivel de $\alpha = 0,01$. Para la segunda muestra la cantidad de ítems con un importante grado de desajuste fue de sólo 6, pero en ninguno de ellos el desajuste fue considerable a nivel de $\alpha = 0,01$.

Se efectuó un análisis de residuos en todo el universo usando el programa RESID, el cual reveló la existencia de sólo 10 ítems con un significativo grado de desajuste entre los datos y el modelo, tres de ellos a nivel de $\alpha = 0,01$. No obstante, de acuerdo con el programa BILOGW, en 6 de dichos ítems el desajuste no era acentuado.

Los residuos estandarizados se distribuyeron de la siguiente manera (véase p. 196 para una explicación):

RE	porcentaje
0 – 1	73,91
1 – 2	23,91
2 – 3	1,88
> 3	0,31

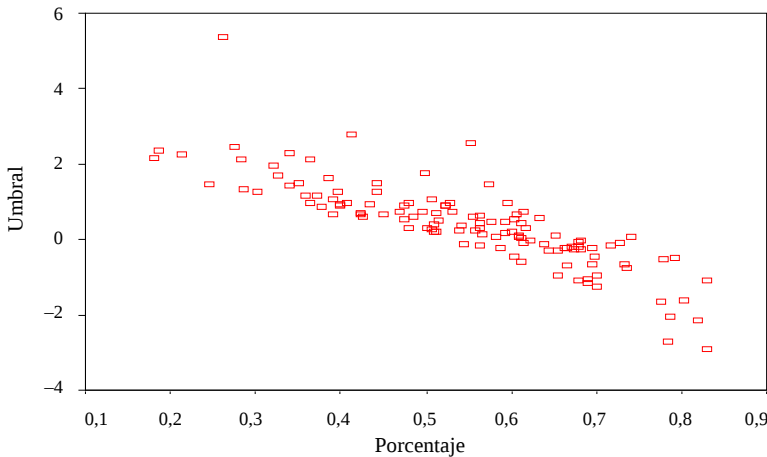
Como la distribución de los RE se acerca mucho a la distribución normal, el análisis de residuos respalda la tesis del ajuste del modelo a los datos.

Comparación entre estadísticas de ítems

La correlación entre valores p de TCM y valores b de IRT fue de $r = -0,84$, y entre la r_{bis} de TCM y los valores a de IRT la correlación fue de $r = 0,63$. La correlación entre los puntajes de tests de TCM y las estimaciones de habilidad de IRT fue de $r = 0,96$. En la Figura 4 se grafican los valores b contra los valores p .

En esta figura parece haber una relación curvilínea entre valores b y p . Puesto que los valores p expresan en realidad la dificultad de los ítems en una escala ordinal, se efectuó una transformación a una escala de intervalos por normalización (véase, por ejemplo, Aiken, 1991, p. 96). Sin embargo, esta normalización no hizo más que aumentar la correlación entre valores b y p a $r = 0,85$.

FIGURA 4: GRÁFICO DE DISPERSIÓN DE VALORES b DE IRT CONTRA VALORES p DE TCM



Con el objeto de investigar más a fondo la invarianza de los parámetros se tomaron 2 muestras aleatorias distintas, de 1.000 individuos cada una, en el universo de personas que rindieron la prueba. Para cada muestra se calcularon las estadísticas de ítems de TCM y se estimaron los parámetros de ítems de IRT. Se realizaron comparaciones entre valores *p* y correlaciones biserials de ambas muestras y entre valores *b* y valores *a*. También se realizaron comparaciones con los mismos valores tomados del universo. El resultado de estas comparaciones aparece en la Tabla 7.

TABLA 7: ESTABILIDAD DE LAS ESTADÍSTICAS DE ÍTEMS DENTRO DE AMBOS MARCOS DE MEDICIÓN. CORRELACIONES ENTRE DIFICULTADES DE ÍTEMS Y DISCRIMINACIONES DE ÍTEMS BASADAS EN LA TCM Y EN IRT

Correlación entre	Universo–muestra A	Universo–muestra B	Muestra A – muestra B
valores <i>p</i>	0,995	0,996	0,989
valores <i>b</i>	0,940	0,946	0,960
<i>r</i> _{bis}	0,931	0,957	0,882
valores <i>a</i>	0,764	0,775	0,683

3.1. Análisis

Los intentos por ajustar un modelo logístico IRT de 3 parámetros al test SweSAT rendido en el último trimestre de 2002 no fueron muy fructíferos. Es difícil obtener una respuesta inequívoca en cuanto al ajuste del modelo a los datos, ya que no existen criterios objetivos. Con todo, el análisis factorial a nivel de ítems fue desalentador, lo mismo que el test BILOGW chi-cuadrado aplicado al universo total de alumnos que rindieron la prueba. Por otra parte, el test RESID resultó alentador, al igual que los tests BILOGW chi-cuadrado aplicados a dos muestras de 1.000 examinados cada una. Sin embargo, los parámetros de ítems no fueron invariantes, y de acuerdo con Hambleton y otros (1991) “...la invarianza y el ajuste del modelo a los datos son conceptos equivalentes” (p. 24). Por tanto, la conclusión debe ser que el modelo logístico de 3 parámetros no se ajustó a los datos del test SweSAT.

Por otra parte, los resultados de los análisis de TCM fueron muy estimulantes. Los índices de dificultad de los ítems, lo mismo que sus índices de discriminación, fueron muy estables entre ambas muestras aleatorias, como también entre el universo y las dos muestras.

3.2. A modo de conclusión

Puesto que el SweSAT ha sido elaborado dentro del marco de la TCM, ha parecido razonable comparar los resultados de los análisis efectuados dentro de IRT con los resultados correspondientes obtenidos en el contexto de la TCM. Las comparaciones empíricas entre los resultados derivados de ambas teorías no son muy habituales. Lawson (1991) comparó el modelo de un parámetro y la TCM en tres series de datos y encontró "...notables similitudes ente los resultados obtenidos a través de métodos de medición clásicos y los resultados obtenidos a través de métodos de rasgo latente³ de un parámetro" (p. 163). Fan (1998) examinó las estadísticas de personas e ítems derivadas de IRT y la TCM en una base de datos de evaluación en gran escala de nivel estatal. Sus resultados indican que las estadísticas de personas e ítems derivadas de ambos marcos de medición eran bastante comparables. También descubrió que la invarianza de las estadísticas de ítems en todas las muestras, característica donde por lo general se considera que radica la superioridad de los modelos de IRT, parecía ser similar dentro de ambos marcos de medición. Tanto Lawson (1991) como Fan (1998) concluyen sus artículos citando pasajes de un discurso de apertura pronunciado por Thorndike en 1982 en una conferencia australiana centrada en modelos de IRT:

Dudo que vaya a observarse un cambio significativo en la mayor parte de los tests, tanto en los elaborados a nivel local como en los estandarizados. Los ítems que escogeremos para un test no diferirán mucho de los que habríamos seleccionado con procedimientos anteriores, y el test resultante seguirá conservando en gran medida las mismas propiedades (p. 12).

En los estudios citados en este trabajo los índices de ítems de TCM eran no sólo comparables con los parámetros de ítems de IRT, sino que además eran por lo general más invariantes entre distintas muestras de examinados. Una posible explicación para estos resultados es que el modelo de IRT no se ajustaba a los datos del test. Pero incluso si los resultados se deben a un deficiente ajuste entre modelo y los datos, la única conclusión razonable es que para los datos de la SweSAT la TCM parece funcionar mejor que IRT.

La SweSAT tiene buena acogida entre los examinados y las universidades de Suecia. La exigencia más importante que se le formula es que

³ Los modelos de IRT se denominan en ocasiones modelos de rasgo latente.

clasifique a los examinados de la manera más justa posible con respecto a su éxito previsto en los estudios. Otros requisitos con que debe cumplir el test son los siguientes:

- Debe estar en concordancia con los objetivos y el contenido de la educación superior.
- No debe tener efectos negativos en la educación secundaria de ciclo superior.
- Debe ser posible puntuar la prueba con rapidez, a bajo costo y objetivamente.
- No debe ofrecer la ocasión para que un individuo mejore sus resultados por medio de ejercicios mecánicos o asimilando principios especiales de resolución de problemas.
- Los examinados deben considerar el test como una experiencia significativa y apropiada.
- Es preciso observar la exigencia de que el reclutamiento esté libre de sesgos. Ningún grupo debe ser discriminado debido a razones tales como género, clase social, etc.

Vale la pena empeñarse en introducir todas las modificaciones que permitan mejorar el test en cualquiera de sus aspectos. Ello requeriría, sin embargo, efectuar cambios que aumentarían el grado de validez de dicho test. Ese debe ser con certeza el principal objetivo de todo cambio: aumentar la validez del test y no hacer que se ajuste a una determinada teoría de medición.

La IRT consiste en una familia de modelos que, según se ha sostenido, son útiles en el diseño, la construcción y la evaluación de tests educacionales. Es de esperar que a medida que se profundicen las investigaciones se vayan resolviendo los restantes problemas técnicos asociados a la aplicación de los modelos. También se espera que en los años venideros se desarrollen modelos más recientes y más aplicables, permitiendo que IRT proporcione soluciones aun mejores a importantes problemas de medición (Hambleton y otros, 1991). Como se mencionó anteriormente, otro importante supuesto de los modelos de IRT es la unidimensionalidad, lo cual significa que los ítems de un test miden una sola habilidad. Hay modelos en los que se supone que se requiere más de una habilidad para explicar el desempeño en un test. Estos así llamados modelos multidimensionales son, a pesar de todo, más complejos y no han sido tan bien desarrollados como los modelos unidimensionales.

Aun cuando en la actualidad IRT no parece ser aplicable a la construcción y el diseño de la SweSAT, se encuentran en marcha algunos trabajos

destinados a investigar si IRT puede emplearse para ecualizar las distintas versiones del test. Emons (1998) realizó un exhaustivo estudio: “Nonequivalent groups IRT observed score equating. Its applicability and appropriateness for the Swedish Scholastic Aptitude Test” (“Ecuación de puntajes observados en IRT para grupos no equivalentes. Su aplicabilidad y conveniencia para la Prueba de Admisión a las Universidades en Suecia”), en el cual utiliza preguntas que habían sido incluidas como ítems ancla en pretests aplicados en el segundo trimestre de 1996, para así ecualizar el test del segundo trimestre de 1997. Puede que en ese estudio, sin embargo, la cantidad de ítems ancla haya sido demasiado exigua para constituir un vínculo que permitiera la ecualización. En algunos subtests el vínculo consistió de sólo 2 a 3 ítems. Estudios similares se realizan continuamente, y en la actualidad los resultados obtenidos con el método tradicional de “ecualización equipercantil de grupos equivalentes” se comparan permanentemente con la ecualización de la IRT empleando vínculos de diferentes magnitudes.

Otra área de estudios sobre SweSAT en que se emplean hoy en día modelos de IRT es el funcionamiento diferencial de los ítems (FDI). Las curvas características del ítem ilustran de un modo muy apropiado el problema del FDI, ya que las curvas muestran la probabilidad de responder correctamente un ítem, dado un cierto nivel de habilidad. La comparación de las curvas para diversos grupos de examinados (en el caso de SweSAT, principalmente hombres y mujeres) corresponde exactamente a la definición más aceptada de FDI. Para estos tipos de estudios puede usarse el puntaje del subtest, y a nivel de subtests la unidimensionalidad parece ser un supuesto aceptable.

Por último, en el caso de los tests adaptativos computarizados (TAC) o tests a la medida, IRT proporciona el único marco teórico adecuado. En los TAC, los ítems que le toca responder a un examinado dependen de su desempeño en los anteriores ítems en el test. Sólo los ítems que entregan más información sobre el examinado son los que se aplican. Los examinados con un alto nivel de habilidad no necesitan responder los ítems más fáciles, y a su vez los examinados con bajo nivel de habilidad no necesitan contestar los ítems de alta dificultad. De esta manera el test puede acortarse considerablemente y aun así entregar la misma información y la misma precisión en la medición que el test convencional más largo. Si en el futuro la SweSAT —o alguna versión de la misma— llegara a transformarse en un TAC entonces resultará imperioso aplicar IRT.

REFERENCIAS BIBLIOGRÁFICAS

- Aiken, L. R. *Psychological Testing and Assessment*. Massachusetts: Allyn & Bacon, 1991.
- Crocker, L. L. y J. Algina. *Introduction to Classical and Modern Test Theory*. New York: Holt, Rinehart & Winston, 1986.
- Emons, W., H., M. (1998). "Nonequivalent Groups IRT Observed Score Equating. Its Applicability and Appropriateness for the Swedish Scholastic Aptitude Test". *Educational Measurement* N° 32 (1998), Umeå, Umeå University, Department of Educational Measurement.
- Fan, X. "Item Response Theory and Classical Test Theory: An Empirical Comparison of their Item/Person Statistics". *Educational and Psychological Measurement*, Vol. 58 N° 3 (1998), p. 357-381.
- Hambleton, R. K. y R. J. Rovinelli. "Assessing the Dimensionality of a Set of Test Items". *Applied Psychological Measurement*, 10 (1986), pp. 287-302.
- Hambleton, R. K.; H. Swaminathan, y H. J. Rogers. *Fundamentals of Item Response Theory*. Newbury: Sage, 1991.
- Hambleton, R. K. y R. W. Jones. "Comparison of Classical Test Theory and Item Response Theory and their Applications to Test Development". *Educational Measurement: Issues and Practice*, 12 (3) (1993), pp. 535-556.
- Hays, W. L. *Statistics*. Londres: Holt, Rinehart and Winston, 1969.
- Henrysson, S. "Gathering, Analyzing, and Using Data on Test Items". En R. L. Thorndike (ed.), *Educational Measurement*. Washington DC: American Council on Education, segunda edición, 1971.
- Lawson, S. *One-Parameter Latent Trait Measurement: Do the Results Justify the Effort?* The Annual Series of the Southwest Educational Research Association, Vol. 1 (1991), pp. 159-168.
- Lord, F. M. *Applications of Item Response Theory to Practical Testing Problems*. Hillsdale NJ: Lawrence Erlbaum, 1980.
- Mislevy, R. L. y R. D. Bock BILOG-W. Manual for BILOG-W. Item Analysis and test Scoring with Binary Logistic Models. Chicago: Scientific Software International, 1990.
- Rogers, J. RESID. Assessment of Fit for Unidimensional IRT Models. Programa desarrollado en University of Massachusetts School of Education, 1994.
- Stage, C. "An Attempt to Fit IRT Models to the DS Sub-test in The SweSAT". *Educational Measurement* N° 19 (1996), Umeå, Umeå University, Department of Educational Measurement.
- Stage, C. "The Applicability of Item Response Models to the SweSAT. A Study of the DTM Sub-test". *Educational Measurement* N° 21 (1997a), Umeå: Umeå University, Department of Educational Measurement.
- Stage, C. "The Applicability of Item Response Models to the SweSAT. A Study of the ERC Sub-test". *Educational Measurement* N° 24 (1997b), Umeå: Umeå University, Department of Educational Measurement.
- Stage, C. "The Applicability of Item Response Models to the SweSAT. A Study of the READ Sub-test". *Educational Measurement* N° 25 (1997c), Umeå, Umeå University, Department of Educational Measurement.
- Stage, C. "The Applicability of Item Response Models to the SweSAT. A study of the WORD Sub-test". *Educational Measurement* N° 26 (1997d), Umeå, Umeå University, Department of Educational Measurement.

- Stage, C. "A Comparison Between Item Analysis Based on Item Response Theory and Classical Test Theory. A Study of the SweSAT Sub-test WORD". *Educational Measurement* N° 29 (1998a), Umeå, Umeå University, Department of Educational Measurement.
- Stage, C. "A Comparison Between Item Analysis Based on Item Response Theory and Classical Test Theory. A Study of the SweSAT Sub-test ERC". *Educational Measurement* N° 30 (1998b), Umeå, Umeå University, Department of Educational Measurement.
- Stage, C. "A Comparison Between Item Analysis Based on Item Response Theory and Classical Test Theory. A Study of the SweSAT Sub-test READ". *Educational Measurement* N° 33 (1999), Umeå, Umeå University, Department of Educational Measurement.
- Thorndike, R. L. (1982). "Educational Measurement: Theory and Practice". En D. Spearritt (ed.), *The Improvement of Measurement in Education and Psychology: Contributions of Latent Trait Theory*. Princeton, NJ: ERIC Document Reproduction Service N° ED 222545. □

LA TRANSFORMACIÓN DE LA VISIBILIDAD*

John B. Thompson

A la luz del enfoque sociológico de la “interacción de los medios de comunicación”, John B. Thompson se refiere en estas páginas al impacto de los medios modernos en la vida social y política. Este enfoque se centra, básicamente, en las nuevas formas de interacción a que dan lugar los medios de comunicación modernos. Estas nuevas formas se caracterizan por no estar sujetas a las coordenadas de tiempo y espacio propias de la comunicación cara a cara. Una de las consecuencias de estas nuevas formas de interacción social tiene que ver con la esfera de lo público y su visibilidad. Los medios modernos, sostiene el autor, han dado origen a un tipo “mediático” de visibilidad, que ha tenido interesantes consecuencias. En la política, por ejemplo, esa visibilidad mediática permite que los políticos alcancen grandes audiencias y manejen su imagen ante ellas, pero, por otro lado, los medios modernos someten a esos mismos actores a niveles de exposición y control inéditos, lo que puede hacer más frágil su posición como titulares del poder. Para ejemplificar estas reflexiones, Thompson echa mano al fenómeno del escándalo políti-

JOHN B. THOMPSON. Profesor de sociología y miembro del Jesus College en la Universidad de Cambridge, Inglaterra. Ha publicado numerosos artículos y libros sobre hermenéutica y teoría del lenguaje, ideología y cultura moderna, teoría social de los medios de comunicación, esfera pública y escándalos políticos. Entre sus publicaciones destacan: *Ideology and Modern Culture: Critical Social Theory in the Era of Mass Communication* (1991), *Los Media y la Modernidad: Una Teoría de los Medios de Comunicación* (1998) y *El Escándalo Político: Poder y Visibilidad en la Era de los Medios de Comunicación* (2001).

* Traducido al castellano por el equipo de traductores de *Estudios Públicos*.

co en la historia moderna. A su juicio, hay tres factores que lo explican: los cambios económicos y tecnológicos experimentados por la prensa, el desarrollo del periodismo como profesión y cambios en la cultura política que harían enfatizar menos el compromiso ideológico que el “carácter” de los actores políticos.

¿Cómo deberíamos entender la naturaleza e impacto de los medios de comunicación? ¿De qué modo la proliferación de medios de comunicación, desde el nacimiento de la imprenta a la internet, altera el carácter de la vida política y social? Obviamente, estas nuevas formas de comunicación producen una enorme expansión de los flujos de información, en cuanto las personas tienen cada vez más acceso a materiales simbólicos que pueden ser transmitidos por medios distintos a la viva voz. Las nuevas formas de comunicación han creado también nuevas oportunidades para la manipulación simbólica, ya que quienes controlan la producción de símbolos pueden usar su poder para moldear el material simbólico de acuerdo con sus objetivos e intereses. Pero también hay modos más profundos en que el desarrollo de los medios de comunicación ha alterado el carácter de la vida política y social. Antes que con el notable aumento de información disponible o con la manipulación ideológica de los materiales simbólicos, estos modos tienen que ver con el cambio en las formas de acción e interacción que caracterizan a las sociedades en que los medios de comunicación se han transformado en una de sus principales características.

En este ensayo delinearé una perspectiva de los medios de comunicación que he desarrollado en varias publicaciones¹. A esta perspectiva la llamo “teoría de la interacción de los medios de comunicación”. Ella es básicamente una concepción sociológica de los medios, en el sentido de que éstos son analizados a la luz de la interacción social que posibilitan y de la que ellos mismos participan. El uso de los medios de comunicación no sólo comprende la transmisión de información a individuos cuyas relaciones recíprocas permanecen inalteradas. Más bien, mediante su uso los individuos crean nuevas formas de acción e interacción que difieren en determinados aspectos de la interacción cara a cara que caracteriza a la mayor parte de nuestros encuentros en la vida cotidiana. El análisis de estas formas de interacción nos provee de una perspectiva sobre las formas en que los medios de comunicación han alterado las condiciones de la vida social y política.

¹ Véase especialmente John B. Thompson, *The Media and Modernity: A Social Theory of the Media* (1995) [versión en castellano por editorial Paidós, 1998].

Una manera fundamental en la que los medios han alterado dichas condiciones tiene que ver con lo que describiré como la transformación de la visibilidad. En las primeras formas de sociedad, la visibilidad de los individuos, acciones y eventos estaba intrínsecamente unida al hecho de compartir un lugar común: los individuos eran visibles en la medida en que podían ser vistos por otros que compartían el mismo contexto espacio-temporal. Con el desarrollo de los medios de comunicación, sin embargo, la visibilidad de los individuos, acciones y eventos es independiente de la circunstancia de compartir un lugar común: los individuos se hacen visibles a otros que se encuentran en contextos distintos y distantes. Las características de estas nuevas formas de visibilidad mediática difieren del tipo de visibilidad que se da en encuentros cara a cara. Y el desarrollo de estas formas de visibilidad mediática da lugar a un conjunto de nuevos fenómenos relativos a los modos y el grado en que la información, imágenes y otros materiales se hacen visibles a los demás. La administración de la visibilidad —un arte esencialmente imperfecto— se ha transformado en una característica inevitable de la política moderna. Pero no importa cuánto se esfuercen los políticos por administrar su visibilidad, ellos no pueden controlarla totalmente. Más adelante intentaré concentrarme en el fenómeno específico de los escándalos políticos, que ilustran bien este carácter incontrolable de la visibilidad mediática.

Los medios de comunicación social y la interacción social

Un supuesto básico de la teoría de la interacción de los medios de comunicación consiste en que los medios deben ser examinados en relación con las formas de acción e interacción que ellos hacen posible y de las cuales, en aquellas ocasiones en que los medios se utilizan, ellos participan. Los medios no son sólo instrumentos técnicos. Son, además, instrumentos para guardar e intercambiar información y materiales simbólicos entre las personas, por lo que están intrínsecamente vinculados a formas de acción e interacción que surgen cuando ellos son utilizados. ¿Cuáles son estas formas de acción e interacción? Podemos empezar a analizarlas distinguiendo entre varios tipos básicos. Un tipo es la interacción que podemos llamar interacción cara a cara. Aquí los participantes se sitúan en una presencia recíproca, comparten un mismo marco espacio-temporal. En otras palabras, la interacción ocurre en un contexto de co-presencia. La interacción cara a cara tiene un carácter “dialógico”, ya que generalmente envuelve un flujo de información y comunicación en dos sentidos: un individuo habla a otro (u otros) y el interpelado puede responder (en principio, al menos), y de

esta manera se desenvuelve el diálogo. Otra característica de la interacción cara a cara es que envuelve una multiplicidad de claves simbólicas: a fin de enviar mensajes e interpretar los mensajes del otro, las palabras pueden ser reemplazadas por gestos, expresiones faciales, cambios en la entonación de la voz, etc.

El funcionamiento de los medios de comunicación da origen a varias formas de interacción mediática, que se diferencian en diversos aspectos de la interacción cara a cara, por lo pronto en sus características espacio-temporales. Mientras la interacción cara a cara siempre ocurre en un contexto de copresencia, la interacción mediática se “extiende” por el espacio y puede también extenderse o comprimirse en el tiempo. En la interacción cara a cara los productores y receptores de mensajes comparten un mismo marco espacio-temporal, pero en la interacción mediática los productores y receptores de mensajes están por lo general separados espacialmente (y pueden o no estarlo también en el tiempo). A través de los medios de comunicación podemos interactuar con otros que no comparten un mismo marco espacio-temporal, y la naturaleza de nuestra interacción está determinada por esta característica y por las propiedades del medio empleado.

Actividades tales como escribir cartas o hablar por teléfono ejemplifican un tipo de interacción mediática. Estas actividades envuelven el uso de un medio técnico para transmitir información o contenidos simbólicos a personas distantes en el espacio o en el tiempo, o en ambas dimensiones a la vez. Debido a que los participantes no comparten el mismo marco espacio-temporal, tienen siempre que considerar cuánta información contextual debe ser incluida en el intercambio, como, por ejemplo, señalando el lugar y fecha al comienzo de una carta o identificándose al iniciar una conversación telefónica. Este tipo de interacción mediática también envuelve una disminución del rango de claves simbólicas, ya que en el contexto de una interacción cara a cara generalmente hay una mayor cantidad y variedad de esas claves.

Un tipo distinto de interacción mediática es aquella generada por la producción y recepción de materiales tales como libros, diarios, programas de radio y televisión, etc.; es decir, los así llamados medios masivos. La denomino “cuasi interacción mediática”. Al igual que los otros tipos de interacción mediática, ésta se prolonga a través del tiempo y espacio, y presenta una cierta disminución del rango de claves simbólicas. Sin embargo, tiene dos características que la diferencian de otras formas de interacción mediática. Primero, en el caso de la cuasi interacción mediática, las formas simbólicas son producidas para un conjunto indefinido de potenciales receptores. En una conversación telefónica, las palabras están dirigidas

a una persona determinada, pero un programa de televisión o un periódico están dirigidos a cualquiera que tenga los medios (culturales y materiales) para recibirlos. Segundo, mientras la forma de interacción comprendida en la escritura de una carta o en una conversación telefónica tiene un carácter generalmente dialógico, la cuasi interacción mediática es predominantemente monológica, es decir, el flujo de la comunicación se desplaza básicamente en un sentido. El lector de un libro o un televidente son receptores de una forma simbólica cuyo productor no requiere (y generalmente no recibe) una respuesta inmediata y directa.

Debido a que la cuasi interacción mediática tiene un carácter monológico y comprende la producción de formas simbólicas para un conjunto indefinido de receptores potenciales, se la entiende mejor como cuasi interacción. No tiene el grado de reciprocidad ni la particularidad interpersonal de otras formas de interacción, sea ésta mediática o cara a cara. La cuasi interacción mediática es, con todo, una forma de interacción. Ella genera una situación social en la que los individuos están vinculados mediante un proceso de comunicación e intercambio simbólico. Ella también genera formas específicas de relaciones interpersonales, vínculos sociales e intimidad (que yo llamo “intimidad no recíproca a distancia”).

El desarrollo de una serie de nuevas tecnologías de comunicación asociadas a los computadores personales e internet puede ser analizado de una manera muy similar. Estas tecnologías crean novedosas situaciones de interacción con características distintivas. Podríamos calificar gruesamente a estas situaciones como variados tipos de “interacción mediática a través del computador”, pero, de hecho, es necesario distinguir entre varios tipos diferentes. El uso del email, por ejemplo, se parece en ciertos aspectos al tipo de interacción mediática comprometida en la escritura de una carta, pero difiere en aspectos importantes. Al igual que la escritura de una carta, el email tiene un carácter dialógico, dirigido a un tercero específico y utiliza la palabra escrita en vez de la oral. Sin embargo, la compresión temporal es mucho mayor y las instituciones mediáticas y condiciones de uso (incluyendo los recursos culturales y materiales) son completamente diferentes. Otras formas de comunicación a través de internet, como grupos de noticias, conversaciones, diarios murales, entre otros, manifiestan características diferentes. Muchas de estas formas de comunicación son más abiertas que el email y de libre acceso para cualquier persona con las habilidades y equipamiento necesarios (a pesar de que algunos pueden tener requisitos especiales de acceso, como claves y/o el pago de cuota de entrada). Pero, a diferencia de la cuasi interacción mediática, también pueden

tener rasgos dialógicos específicos, tales como el chat a tiempo real, ordenar y/o bajar material de sitios de internet o contactarse vía email².

Si analizamos los medios de comunicación en base a la teoría de la interacción, podemos obtener una perspectiva sociológica distintiva de su desarrollo e importancia histórica. Previo al auge de la imprenta en la Europa de la Edad Media tardía y de la era moderna temprana (y hasta hace poco en otros lugares del mundo), el intercambio de información y contenido simbólico era, para la mayoría de las personas, un proceso que se llevaba a cabo exclusivamente en el contexto de la interacción cara a cara. Había varias formas de interacción y cuasi interacción mediática, pero estaban restringidas a un sector relativamente pequeño de la población. Mas, con el surgimiento de la industria de la imprenta en los siglos XV y XVI y el desarrollo de varios tipos de medios de comunicación electrónicos en los siglos XIX y XX, los “componentes” de la interacción social han cambiado. La interacción cara a cara no ha sido desplazada, pero sí complementada por otras formas de interacción que han asumido un rol crecientemente importante. Cada vez más, los individuos pueden adquirir información y contenido simbólico de fuentes distintas a las personas con las que interactúan directamente en su vida cotidiana, y acceder a “conocimiento no local”, el que pueden incorporar reflexivamente en sus procesos autoformativos. La creación y renovación de las tradiciones son procesos que han ido quedando progresivamente vinculados al intercambio simbólico mediático; las tradiciones no han desaparecido con el desarrollo de las sociedades modernas, pero gradualmente han perdido su anclaje en los emplazamientos de la vida cotidiana. El desarrollo de los medios de comunicación crea nuevos campos de acción e interacción que involucran formas diferentes de visibilidad, en las cuales las relaciones de poder pueden cambiar de modo rápido, dramático e impredecible. Examinemos este último punto en más detalle.

La transformación de la visibilidad

Una consecuencia del surgimiento de la interacción y cuasi interacción mediática fue la transformación gradual de lo que podríamos llamar “lo público”^{*} y, muy ligada a ella, una transformación en las formas en que

² Para una discusión sobre comunicación a través de internet desde la perspectiva de la teoría interaccional, véase James Slevin, *The Internet and Society* (2000).

^{*} *Publicness*, en el original. Traducido aquí y más adelante como “lo público” o “publicidad”, esta última entre comillas. (N. del T.)

los individuos y los acontecimientos se hacen “visibles” a otros. La distinción entre lo público y lo privado tiene una larga historia en el pensamiento social y político occidental³. Se remonta a los debates filosóficos sobre la vida de la polis en la Grecia clásica y al temprano desarrollo del derecho romano, con su concepción de *res publica*. Sin embargo, en la Edad Media tardía y en la época moderna temprana, la distinción público/privado comenzó a adquirir nuevos significados, debido, en parte, a las transformaciones institucionales que se daban en la época. Si nos centramos en el desarrollo de las sociedades occidentales desde el período medieval tardío, podremos distinguir dos sentidos diferentes de la dicotomía público/privado. Estos dos sentidos no agotan los significados de dichos términos, pero sí capturan algunas de las formas más importantes en las cuales se usan.

El primer sentido de esta dicotomía tiene que ver con la relación entre el ámbito del poder político institucionalizado, conferido crecientemente al Estado soberano, y los ámbitos de la actividad económica y las relaciones personales que quedaban fuera del control político directo. De este modo, desde mediados del siglo XVI en adelante, el término “público” comenzó a significar, progresivamente, actividad o autoridad relacionada con el Estado o derivada de éste, mientras que “privado” se refería a aquellas actividades o esferas de la vida que estaban excluidas o separadas de él. Sin embargo, podemos distinguir un segundo sentido de la dicotomía público/privado. De acuerdo con éste, “público” significa “abierto” o “disponible al público”. Lo que es público, en este sentido, es aquello visible u observable, lo efectuado ante espectadores, lo abierto a todos o muchos para ver u oír. Lo privado, en contraste, es lo oculto a la vista, lo dicho o hecho en privado, confidencialmente o al interior de un círculo restringido de personas. En este sentido, la distinción entre público y privado se refiere a publicidad versus privacidad, a apertura versus confidencialidad, a visibilidad versus invisibilidad. Un acto público es un acto visible, efectuado abiertamente para que cualquiera pueda verlo; un acto privado es invisible, un acto efectuado confidencialmente y tras puertas cerradas.

Aquí quiero concentrarme en el segundo sentido de esta dicotomía y usarlo como contrapunto para explorar el impacto de los medios de comunicación social. Previo al desarrollo de los medios de comunicación, la “publicidad” o visibilidad de un individuo o evento estaba vinculado al hecho de compartir un lugar. Un evento se transformaba en acontecimiento

³ Véase especialmente a Jürgen Habermas, *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society* (1989), y Norberto Bobbio, *Democracy and Dictatorship: The Nature and Limits of State Power* (1989).

público cuando se realizaba frente a una pluralidad de individuos que estaban físicamente presentes —como en el caso, por ejemplo, de una ejecución en la Europa medieval que se llevaba a cabo ante un grupo de espectadores reunidos en la plaza pública—. Quiero describir esto como la “‘publicidad’ tradicional de la copresencia”. Esta “publicidad” tradicional se valía de la riqueza de las claves simbólicas que son características de la interacción cara a cara. Ella involucraba tanto la vista como el sonido, aparición visual como palabra oral: el evento público era un espectáculo que, para aquellos relativamente escasos individuos que estaban presentes cuando ocurría, podía ser visto, oído o, quizás, hasta sentido de alguna manera. Más aún, ya que la “publicidad” de la copresencia implica la reunión de individuos en un lugar común, ella tiene un carácter potencialmente dialógico. Los individuos que hablan o actúan en ocasiones de ese tipo, lo hacen ante otros que pueden, en principio, contribuir al evento, ya sea hablando o desplegando otras formas de conductas propias de los espectadores (abucheando, silbando, aplaudiendo, vitoreando, etc.), aun cuando en la práctica no lo hagan.

El desarrollo de los medios de comunicación creó nuevas formas de lo público, muy diferentes de la “publicidad” tradicional de la copresencia. El factor clave de estas nuevas formas es que, con la extensión de la accesibilidad que permiten los medios de comunicación, la “publicidad” de los individuos, acciones o eventos ya no está ligada a la circunstancia de compartir un lugar común. Una acción o evento puede hacerse público al ser grabado y transmitido a otros que no estén físicamente presentes en el momento y lugar de su ocurrencia. Las acciones o eventos pueden adquirir una “publicidad” —que llamaré “‘publicidad’ mediática”— independiente de su capacidad de ser vistos u oídos directamente por una pluralidad de individuos copresentes. Por supuesto que estas nuevas formas de “publicidad” mediática no desplazaron completamente a la “publicidad” tradicional de la copresencia. La forma tradicional continúa jugando un papel importante en las sociedades modernas; prueba de ello es la permanente realización de reuniones públicas, concentraciones de masas, debates políticos en los escenarios cara a cara de los parlamentos y otros órganos de decisión, y así sucesivamente. En cuanto los nuevos medios de comunicación se hicieron más ubicuos, las nuevas formas de “publicidad” comenzaron a complementar, extender, transformar y, en algunos casos, desplazaron a la forma tradicional de lo público.

Podemos rastrear brevemente algunos aspectos de este proceso, considerando ciertos episodios de la extensa aunque desatendida historia de las relaciones entre las cambiantes formas de “publicidad” y el ejercicio

del poder político. Previo al desarrollo de la imprenta y otros medios de comunicación, la “publicidad” de los gobernantes derivaba principalmente de su aparición física ante otros en el contexto de la copresencia. La mayoría de las veces estas apariciones se restringían a los círculos relativamente cerrados de la asamblea o la corte: la visibilidad requería la copresencia y los gobernantes solían ser visibles sólo para aquéllos con los que interactuaban rutinariamente en los escenarios del cara a cara de la vida cotidiana. Su audiencia estaba conformada principalmente por miembros de la elite gobernante o por individuos que participaban en la vida social de la corte.

En ocasiones los gobernantes aparecían ante concurrencias más vastas, que incluían, entre otros, a algunos de los súbditos. Estas ocasiones podían ser eventos públicos mayores como coronaciones, funerales reales o marchas victoriosas. La pompa y ceremonia de tales eventos permitían al gobernante mantenerse a cierta distancia de sus súbditos, y al mismo tiempo permitían que éstos pudiesen ver y celebrar temporalmente a sus gobernantes en un contexto de copresencia. Sin embargo, en las sociedades antiguas o medievales, la mayoría de los individuos tenían muy pocas ocasiones, si alguna, de ver a los gobernantes más poderosos. Quienes vivían en áreas rurales o en las regiones periféricas de un imperio o reino muy rara vez tenían la posibilidad de ver en carne y hueso al emperador o al rey. Aparte de las marchas reales —transitorias y relativamente poco frecuentes—, la mayoría de las apariciones públicas del monarca se daban en el centro político, en los salones de palacio o en las calles y plazas de la ciudad capital⁴.

Sin embargo, con el desarrollo de la imprenta y otros medios de comunicación los gobernantes comenzaron a adquirir un tipo de “publicidad” desligada de su aparición física ante un conglomerado de personas. Los gobernantes usaron los nuevos medios de comunicación no sólo para promulgar decretos oficiales, sino también para fabricarse una imagen personal que podía ser transmitida a otros en lugares distantes. Los monarcas de la temprana Europa moderna, tales como Luis XIV de Francia o Felipe IV de España, eran muy versados en las artes de la creación de imagen⁵. Sus imágenes eran fabricadas y exaltadas no sólo en los medios de comunicación tradicionales, tales como la pintura, el bronce, la piedra o la tapicería, sino también en el novedoso medio de la imprenta, incluyendo grabados, aguafuertes, panfletos y periódicos. Bajo el reinado de Luis XIV, por

⁴ Véase Clifford Geertz, “Centers, Kings, and Carisma: Reflections on the Symbolics of Power” (1983), pp. 121-46.

⁵ Véanse Peter Burke, *The Fabrication of Louis XIV* (1992) y J. H. Elliott, “Power and Propaganda in the Spain of Phillip IV” (1985), pp. 145-73.

ejemplo, periódicos como la *Gazzete de France*, que aparecía dos veces por semana, y el *Mercure Gallant*, publicado mensualmente, dedicaban espacios regulares a los actos del rey. Pese a que las monarquías de la temprana Europa moderna eran principalmente sociedades cortesanas, en las que los gobernantes interactuaban con las elites en los palacios y otros lugares especiales, las imágenes de los monarcas y el relato de sus acciones se difundían a través de la imprenta mucho más allá de estos restringidos círculos. La propagación de estas imágenes y relatos hacía que las actividades de los gobernantes fuesen cada vez más visibles a una pluralidad de individuos que no estaban en posición de encontrarse con ellos (o con otros miembros de la elite política) en el curso de sus vidas cotidianas. Gradualmente, la “publicidad” de los gobernantes y de otros individuos, la visibilidad de sus actos, de sus expresiones y, por supuesto, de sí mismos, se fue desligando de su aparición frente a personas que se congregaban en un mismo lugar físico-temporal.

El desarrollo de los medios de comunicación electrónicos como la radio y, sobre todo, la televisión, representó en alguna medida la continuación de un proceso que se había puesto en marcha con el advenimiento de la imprenta, pero en otros aspectos constituyó un nuevo inicio. Tal como en el caso de la imprenta, los medios electrónicos crearon un tipo de “publicidad” desligada de la circunstancia de compartir un mismo lugar físico, una “publicidad” que, con la creciente accesibilidad a medios de comunicación a escala nacional e incluso internacional, repercutió en la vida de una proporción cada vez mayor de la población. No obstante, el tipo de “publicidad” generada gracias a los medios de comunicación electrónicos era diferente, en ciertos aspectos, de aquella creada por la imprenta. Los medios electrónicos hicieron posible que la información y el contenido simbólico se pudiesen transmitir a largas distancias con pequeña o ninguna demora. Desde entonces, los medios electrónicos crearon una forma de “publicidad” caracterizada, al menos en principio, por lo que podríamos llamar “simultaneidad desespacializada”: personas distantes podían ser vistas de manera virtualmente simultánea, podían oírse en el mismo instante en que hablaban o verse en el mismo momento en que actuaban, a pesar de que no compartían el mismo lugar físico con los individuos para los cuales eran visibles. Más aún, los medios de comunicación electrónicos se caracterizan por una riqueza de claves simbólicas que permiten reproducir algunos de los rasgos de la interacción cara a cara, pese a que las propiedades espaciales de la interacción cara a cara y de la cuasi interacción mediática son radicalmente diferentes. La radio permitió codificar la cualidad oral de la voz humana y transmitirla a personas distantes, mientras que la televisión hizo posible

grabar y retransmitir claves visuales y orales. Por lo tanto, con el advenimiento de la televisión, los individuos pueden ver personas, acciones o eventos, así como oír la palabra oral u otros sonidos de una manera tanto simultánea como desespacializada. En la era de la televisión, la “publicidad” mediática se define cada vez más por la visibilidad, en el sentido estricto de visión (la capacidad de ser visto con los ojos), aunque este nuevo campo de visión es completamente diferente de aquel que tienen los individuos en sus encuentros cotidianos con otros.

La manera en que los líderes políticos aparecían ante los demás fue modelada por las cambiantes formas de “publicidad” creadas por los medios de comunicación. Con el surgimiento de la radio, los líderes políticos pudieron hablar directamente a miles, e incluso millones, de personas, de un modo que permitió una particular forma de intimidad, llamada “intimidad no recíproca a distancia”, muy distinta de aquella característica de las relaciones orador-audiencia de las concentraciones multitudinarias. En los días previos al sonido amplificado, un orador que quería dirigirse a una concentración masiva tenía que sacar la voz con gran fuerza; los oradores se paraban sobre tarimas para ser vistos, y frecuentemente usaban un lenguaje encendido que pudiera provocar una respuesta colectiva. Con la llegada de la radio, la distancia retórica cedió paso, en parte, a la intimidad mediática; la encendida oratoria del discurso apasionado podía reemplazarse por la intimidad coloquial de una conversación junto al fuego de la chimenea⁶. Si se añade la riqueza visual de la televisión, el escenario está preparado para el florecimiento de un nuevo tipo de intimidad en la esfera pública. Ahora los líderes políticos pueden dirigirse a sus gobernados como si fueran miembros de su familia o amigos. Y dada la capacidad de la televisión de transmitir tomas de primeros planos, los individuos pueden escrutar las acciones y declaraciones de sus líderes, sus expresiones faciales, su aspecto personal, modales y lenguaje corporal, entre otras cosas, con esa clase de atención reservada alguna vez para aquellos con quienes se compartía una relación íntima.

El desarrollo de nuevos medios de comunicación dio paso, así, a un nuevo tipo de “publicidad” desespacializada que permitió una forma íntima de presentación personal, libre de las restricciones de la copresencia. Estas fueron las condiciones que facilitaron el surgimiento de lo que podríamos llamar sociedad de la “revelación” personal: una sociedad en la cual era posible, y, por cierto, crecientemente común para los líderes políticos y otros individuos, aparecer ante audiencias distantes y dejar al descubierto

⁶ Véase Kathleen Hall Jamieson, *Eloquence in an Electronic Age: The Transformation of Political Speechmaking* (1988), especialmente el Cap. 3.

alguna característica de su persona o de su vida personal. La lejanía impersonal de la mayoría de los líderes políticos en el pasado comenzó a ser reemplazada por una nueva especie de intimidad mediática gracias a la cual los políticos pudieron presentarse no sólo como líderes sino también como seres humanos, como individuos comunes que pueden dirigirse a sus gobernados como conciudadanos, revelando selectivamente aspectos de su vida y su carácter a la manera de una conversación o, incluso, confesión. Lo que se perdió en este proceso fue algo del aura, de la “grandeza” que rodeaba en el pasado a los líderes políticos y a las instituciones, un aura que dependía, en parte, de la lejanía de los líderes y de la distancia a la que se mantenían de los individuos sobre los que gobernaban⁷. Pero se ganó la capacidad de hablar directamente a los gobernados, de aparecer ante ellos como seres de carne y hueso con los que se puede sentir empatía e incluso simpatizar, y de dirigirse a ellos no como a súbditos sino como a amigos. En breve, los líderes políticos adquirieron la capacidad de presentarse como “uno más de nosotros”.

No obstante, la visibilidad mediática es un arma de doble filo. Así como los medios de comunicación han creado nuevas oportunidades para que los líderes políticos aparezcan ante los demás de una forma y en una escala nunca antes vistas, crearon también nuevos riesgos. La visibilidad mediática fue un regalo para los aficionados a modelar su imagen o promover sus fines a través de los medios. Sin embargo, los medios de comunicación no estaban reservados exclusivamente para los líderes políticos. Las formas mediáticas de comunicación podían ser usadas no sólo para promover y exaltar a dichas personalidades, sino también para atacarlas y denunciarlas. Los panfletistas y otros en la temprana Europa moderna solían emplear la palabra impresa para burlarse y caricaturizar a aquellos que ejercían el poder. Luis XIV, por ejemplo, fue comúnmente retratado en la prensa real como un héroe de guerra, pero era satirizado por los panfletistas como un cobarde mujeriego. Es más, ya que los medios tenían la capacidad de hacer visibles espacios de acción antes ocultos a la vista, y ya que crearon un complejo campo de imágenes y un flujo de información muy difíciles de controlar, también dieron origen a nuevas formas de eventos mediáticos que tenían el potencial de trastornar y socavar las cuidadosamente calculadas presentaciones personales de los líderes políticos y otros. Un tipo de evento mediático que ilustra muy bien esto, y que ha adquirido un particular significado en las últimas décadas, es el fenómeno del escándalo político.

⁷ Este punto está descrito muy bien por Joshua Meyrowitz en su obra *No Sense of Place: The Impact of Electronic Media on Social Behavior* (1985), p. 270 en adelante.

El surgimiento del escándalo político

Los orígenes del escándalo político, como evento mediático, se remontan a la cultura panfletaria de los siglos XVII y XVIII⁸. Durante el período de la Guerra Civil inglesa hubo una proliferación de panfletos y escritos antimonárquicos, los que fueron condenados como heréticos, blasfemos, injuriosos y “escandalosos”. En 1645, por ejemplo, Marchamont Nedham, prominente escritor y coeditor de la publicación antimonárquica *Mercurius Britanicus*, lanzó un ataque contra el rey acusándolo de tener “una mente sucia, las manos ensangrentadas y un corazón lleno de promesas quebrantadas”, además de burlarse de su defectuosa oratoria⁹. El ataque fue condenado por los lords como un “oprobio* a la persona del rey” y tanto Nedham como el impresor y representante del *Britanicus* fueron brevemente encarcelados por la ofensa cometida. En Francia, un género particular de literatura política subversiva, que incluía los *libelles* y las *croniques scandaleuses*, emergió a principios del siglo XVIII, la cual pretendía dar cuenta de la vida privada de los reyes y cortesanos y los retrataba desde un ángulo poco elogioso¹⁰. Estos irreverentes relatos con frecuencia se presentaban como auténticas descripciones de lo que ocurría tras la fachada del poder, con supuestos fragmentos de correspondencia secreta y conversaciones privadas.

En el contexto de la cultura panfletaria de los siglos XVII y XVIII, la palabra “escándalo” se usaba principalmente en el sentido de alegaciones blasfemas, injuriosas y difamatorias expresadas en forma impresa. Pese a ello, a fines del siglo XVIII y a comienzos del XIX, el uso de la palabra “escándalo” con relación a los medios de comunicación empezó a cambiar; el término se apartó gradualmente de la idea de blasfemia y sedición, y cada vez más se aplicó a una serie de fenómenos que tenían ciertas características distintivas. Estos fenómenos eran de muy variada índole, pero, en general, suponían la revelación, a través de los medios de comunicación, de algún acto o actividad previamente oculto que involucraba la transgresión de ciertos valores o normas y cuya revelación provocaba una respuesta desaprobatoria. Actividades llevadas a cabo en la clandestinidad o en privado se hicieron súbitamente visibles en el ámbito público y la condena de ellas en la prensa y otros medios de comunicación sirvieron, en parte, para

⁸ Para mayor detalle del argumento enunciado en los párrafos siguientes, véase mi obra *Political Scandals: Power and Visibility in the Media Age* (2000).

* *Scandalous*, en el inglés original. (N. del T.)

⁹ *Mercurius Britanicus*, 92 (4 de agosto de 1645), p. 825.

¹⁰ Véase Robert Darnton, *The Forbidden Best-Sellers of Pre-Revolutionary France* (1996).

que tal evento se considerara como un escándalo. No todos los escándalos se encontraban entrettejidos con formas mediáticas de comunicación, pero muchos sí, y el escándalo, como evento mediático, emergió gradualmente como un género reconocible en el curso del siglo XIX.

Diversas condiciones favorecieron este desarrollo, pero dos fueron particularmente importantes. La primera, el cambio de las bases económicas y tecnológicas de la prensa. Los primeros periódicos de los siglos XVII y XVIII eran, en general, empresas a pequeña escala comercial, a menudo manejadas como negocios familiares. Los tirajes eran relativamente reducidos y los precios algo elevados, y los periódicos estaban dirigidos principalmente a las elites comerciales y políticas. Mas, el desarrollo tecnológico de principios del siglo XIX revolucionó las condiciones de producción, facilitando grandes tirajes a un costo relativamente bajo. Al mismo tiempo, el crecimiento de la población alfabeta creó un mercado en expansión para materiales impresos. Desde 1830, la circulación de algunos periódicos establecidos comenzó a crecer de manera significativa y aparecieron otros nuevos a bajo precio y dirigidos a un público más amplio. Estas publicaciones nuevas y económicas adoptaron un estilo de presentación más ligero y vivaz. Dedicaban un espacio considerable a historias de crímenes, violencia sexual, apuestas y deporte. Solían centrarse en personas individuales y los relatos se ocupaban de algunos aspectos de sus vidas. Convirtieron a lo que hoy sería descrito como “historia de interés humano” en un producto característico de los periódicos.

Un segundo cambio importante fue el surgimiento del periodismo como profesión, a finales del siglo XIX¹¹. Los dueños y editores de los periódicos comienzan a depender cada vez más de escritores y reporteros a quienes se contrataba y pagaba para que reunieran noticias y escribieran historias. A medida que las asociaciones de escritores y reporteros se expandían, comenzó a surgir un ethos profesional que definía los principios de la correcta práctica periodística. Ese ethos enfatizaba, sobre todo, el deber de descubrir los hechos a través de un proceso de pesquisas e investigación. No obstante, también reconoció la necesidad de presentar los hechos de una manera vívida, colorida y atrayente. Objetividad y entretenimiento eran los ideales gemelos de la profesión periodística emergente.

Estos y otros desarrollos fueron las condiciones que dieron forma al surgimiento del escándalo como suceso mediático. Hacia fines del siglo XIX, el escándalo mediático se había convertido en un rasgo relativamente

¹¹ Véase Michael Schudson, *Discovering the News: A Social History of American Newspapers* (1978), Cap. 1.

común del paisaje político en países como Gran Bretaña y Estados Unidos. En Gran Bretaña había una serie de editores como W. T. Stead de la *Pall Mall Gazette*, Henry Labouchere del *Truth* y Ernest Parke del *North London Press*, comprometidos cual cruzados en combatir las injusticias sociales y quienes no temían usar sus periódicos como un vehículo para alcanzar sus fines. Stead estuvo dispuesto, por ejemplo, a apoyar la lucha contra la prostitución y el tráfico de niñas, y en julio de 1885 escribió una serie de artículos en la *Gazette* bajo el título “The Maiden Tribute of Modern Babylon” (“El virginal tributo de la Babilonia Moderna”), donde dejaba al descubierto las duras realidades del tráfico sexual infantil. Los artículos relataron, con detalles gráficos, la historia de una pequeña niña que había sido prostituida y vendida en el extranjero; también atacaron la corrupción policial y la hipocresía de los ricos y poderosos por hacer la vista gorda frente a una tragedia que ayudaban a crear. Estos artículos fueron un éxito y contribuyeron a aumentar la presión para que el Parlamento aprobara el Proyecto de Reforma de la Ley Penal, que aumentaba la edad de consentimiento a los dieciséis años y prohibía el secuestro con propósitos sexuales de niñas menores de dieciocho.

Ernest Parke —del *North London Press*— también se vio involucrado en un gran escándalo en torno a una casa en Cleveland Street, cerca de Tottenham Court Road en el West End de Londres. La casa era un burdel masculino manejado por Charles Hammond, quien empleaba jóvenes muchachos para servir a su clientela. El escándalo salió a la luz por una investigación policial de rutina de un incidente no relacionado. Dos hombres fueron finalmente arrestados y procesados en el Old Bailey en septiembre de 1889; se declararon culpables del cargo de indecencia mayor y fueron sentenciados a cuatro y nueve meses de prisión. Algunos sospecharon que se había llegado a un arreglo: los acusados habrían sido incitados a declararse culpables a cambio de condenas menores, y de esa manera evitar que otros se vieran implicados. Se rumoreaba que la casa de Cleveland Street contaba, entre sus clientes, con varias figuras públicas prominentes, incluidos lord Somerset, el conde de Euston y el príncipe Albert Victor, hijo mayor del Príncipe y la Princesa de Gales y segundo en la línea de sucesión al trono. El 16 de noviembre de 1889 Ernest Parke publica un artículo en el *North London Press* en el que menciona a lord Somerset y al conde de Euston en conexión con el “sumamente detestable escándalo en Cleveland Street” y sugiere que a estas personas se les había permitido escabullirse porque “su procesamiento hubiera revelado el hecho de que un personaje mucho más distinguido y encumbrado estaba involucrado en sus

repugnantes crímenes”¹². Lord Euston demandó exitosamente a Parke por calumnia y salió del escándalo con su reputación, en gran medida, intacta. Lord Somerset, por otro lado, nunca se recuperó. Huyó del país en octubre de 1889, justo antes de que se dictara una orden judicial para su arresto (muchos estaban convencidos de que había sido prevenido) y pasó el resto de su vida exiliado en Francia, donde murió en relativa oscuridad en 1926.

El escándalo de Cleveland Street fue uno de tantos que involucraron a prominentes figuras públicas y que aparecieron en la prensa inglesa a finales del siglo XIX. Otros, incluido el controvertido caso de Oscar Wilde (quien eligió —torpemente como se pudo apreciar— demandar al marqués de Queensberry, acción que hizo pública su homosexualidad y, por consiguiente, selló su destino), la caída de sir Charles Dilke (una naciente estrella del Partido Liberal cuya carrera política quedó irremediabilmente dañada por los eventos que rodearon una demanda de divorcio en la que fue señalado como cómplice) y la ruina de Charles Parnell (el carismático líder del Partido Parlamentario Irlandés, cuya carrera política resultó destruida por revelaciones acerca de su aventura con la señora Katherine O’Shea). Muchos de estos escándalos involucraban sexo, tanto heterosexual como homosexual, y en ese sentido fueron moldeados por el clima moral y legal de la tardía época victoriana (un contexto en el que el divorcio era raro, el adulterio era fuertemente rechazado por muchos y los actos de homosexualidad masculina eran ilegales.) Sin embargo, también hubo escándalos que involucraban acusaciones de fraude y corrupción (a pesar de que en el siglo XIX éstos eran más comunes en Estados Unidos que en Inglaterra). Para fines del siglo XIX, un tipo característico de evento, el escándalo mediático, se convirtió en un rasgo relativamente común de la vida social y política. De este modo, la revelación a través de los medios de comunicación de actividades que previamente se habían mantenido ocultas (o que sólo conocía un pequeño círculo de personas), actividades que eran moralmente criticables, podían acarrear consecuencias muy dañinas para los individuos envueltos y para las organizaciones (incluidos gobiernos y partidos políticos) de las que eran parte.

Así como el siglo XIX fue la cuna del escándalo mediático, el siglo XX vino a ser su verdadero hogar. Una vez que esta forma característica de evento había surgido, se convertiría en un género reconocible que algunos buscarían producir activamente, mientras que otros, especialmente aquéllos prominentes en la vida pública, buscarían evitar, con variados grados de discreción y éxito. El carácter y la frecuencia de los escándalos políticos

¹² *North London Press*, 16 de noviembre de 1889, p. 5.

varían considerablemente de un contexto nacional a otro, dependiendo de una serie de circunstancias políticas y sociales específicas. En países como Inglaterra y Estados Unidos hubo escándalos políticos significativos durante las primeras décadas del siglo XX, tales como el escándalo Marconi en Inglaterra, que estuvo cerca de derribar al gobierno liberal en 1913, y el escándalo *Teapot Dome* en Estados Unidos, que sacudió a la administración Harding en los años veinte. Sin embargo, es indudable que el fenómeno del escándalo político se convirtió en un rasgo particularmente frecuente de la vida pública inglesa y estadounidense, y de una serie de democracias liberales modernas, en el período posterior a los primeros años de la década del sesenta. ¿Por qué? ¿Cómo podemos explicar la creciente frecuencia del escándalo político en las últimas décadas?

El escándalo político hoy

Una manera de responder a esta pregunta sería argumentar que la creciente frecuencia del escándalo político es sintomática de una decadencia en los estándares morales de los líderes políticos, tanto respecto a su conducta personal como a su probidad general en la conducción del cargo. Así, se podría señalar, por ejemplo, que los escándalos sexuales —rasgo común de la vida política inglesa desde el caso Profumo, tal como aquellos que asediaron a la presidencia de Bill Clinton— son reflejo de una declinación general en los estándares morales. De la misma manera, podría argüirse que los escándalos políticos que surgieron en Estados Unidos en los años setenta y ochenta, así como los escándalos político-financieros en la forma de “dinero por preguntas” (*cash for questions*) que destruyeron las carreras de muchos políticos británicos en los años noventa, son efecto de los niveles declinantes de probidad entre los políticos. No obstante esta puede ser una explicación plausible, hay poca evidencia para apoyarla. En efecto, no está del todo claro que, en términos generales, los estándares morales de los políticos de hoy sean significativamente más bajos que aquellos a los que adherían los políticos en el pasado. Las aventuras extramaritales de Kennedy son el más obvio ejemplo: sólo unos pocos presidentes norteamericanos, al parecer, tuvieron amoríos que lograron permanecer como secretos bien guardados en su época¹³. Es probable que la creciente frecuencia de los escándalos políticos tenga menos que ver con una decadencia general en los estándares morales de los líderes políticos que con las

¹³ Véase Wesley O. Hagood, *Presidential Sex: From de Founding Fathers to Bill Clinton* (1998).

cambiantes formas y grado en que sus actividades se hacen visibles en el ámbito público.

Hay una serie de cambios importantes que subyacen bajo la periodicidad cada vez mayor de los escándalos políticos en las décadas recientes; aquí me centraré en tres. El primero son las cambiantes tecnologías de comunicación y vigilancia. El siglo XX ha sido testigo de una verdadera revolución en las tecnologías del registro, procesamiento y transmisión de información y comunicación. Estas nuevas tecnologías, especialmente aquellas asociadas a la televisión, han contribuido a hacer mucho más visibles a los líderes políticos en el ámbito público, y han ayudado a asegurar que sus maneras de aparecer ante los demás presenten un grado de intimidad y revelación personal que raras veces se vio en el pasado. Y los líderes políticos, cuanto más esfuerzos hacen por presentarse a través de los medios de comunicación como individuos comunes con vidas corrientes, mayor es la probabilidad de que las audiencias a las que se dirigen terminen evaluándolos por su carácter individual, sinceridad, honestidad e integridad. Presentándose como personas ordinarias con vidas privadas y compromisos personales, y con sus propios motivos y razones para hacer lo que hacen, los líderes políticos dan al carácter y a la integridad una importancia cada vez mayor en la vida pública. Sin embargo, el carácter es un atributo que puede, de igual modo, condenarlos.

Estas cambiantes tecnologías son importantes por otra razón: hacen cada vez más difícil tejer un velo de privacidad alrededor de la conducta privada de los líderes políticos y otras figuras públicas. Aparatos cada vez más sofisticados, como aquellos relacionados con la grabación secreta de conversaciones, la fotografía a larga distancia, la interceptación encubierta de las telecomunicaciones y la búsqueda y recuperación de comunicación electrónica codificada digitalmente, proveen de un poderoso surtido de equipos que pueden ser usados para incrementar la filtración hacia el ámbito público de conductas secretas, donde, unidas a las acciones de las organizaciones mediáticas y paramediáticas, pueden convertirse en eventos altamente visibles. Sería algo excesivo decir que estas nuevas tecnologías anuncian el “fin de la privacidad”, como algunos estudiosos han sugerido¹⁴. Pese a ello, es indudable que, en parte debido a la creciente accesibilidad de esas tecnologías, las condiciones sociales de la privacidad están cambiando de modo fundamental. Las nuevas tecnologías han creado poderosos medios para “oír a escondidas”. Conversaciones o interacciones que los individuos

¹⁴ Véase Reg Whitaker, *The End of Privacy: How Total Surveillance is Becoming a Reality* (1999).

creen privadas (tanto realizadas cara a cara o con la ayuda de tecnologías uno a uno, como el teléfono) pueden, de hecho, ser interceptadas y grabadas por medios encubiertos, y posteriormente puestas a disposición de muchos miles o millones de terceros a través de los medios de comunicación. Palabras expresadas o acciones llevadas a cabo como comunicación o conductas privadas pueden, de manera inesperada, adquirir un carácter público, haciéndose visibles de modos no anticipables, posiblemente muy vergonzosos y quizás hasta seriamente incriminatorios (tal como Mónica Lewinsky y Bill Clinton, entre muchos otros, lo comprobaron.)

Un segundo factor es el cambio en la cultura del periodismo. Como vimos, el surgimiento del periodismo como profesión proviene de fines del siglo XIX, y había, en ese tiempo y a principios del siglo XX, una serie de periodistas y editores que se consideraban investigadores en busca de develar realidades ocultas para llevarlas a la atención del público, no sólo editores “cruzados” como W. T. Stead en Inglaterra y Joseph Pulitzer en Estados Unidos, sino también los periodistas llamados “cazadores de escándalos” como Lincoln Steffens e Ida Tarbell. Después de este temprano florecimiento, la tradición del periodismo inquisitivo entró en un período de relativa declinación hasta los años sesenta, cuando cobró renovado impulso con los tumultuosos eventos políticos de la década. En Estados Unidos, el movimiento por los derechos civiles y la lucha contra la Guerra de Vietnam fueron, quizás, los más importantes, pero otros, tales como el movimiento feminista y el cuestionamiento a las actitudes tradicionales frente a la sexualidad, fueron también de gran significación. Esta cultura cuestionadora y crítica, por sí sola no transformó las prácticas periodísticas, pero sí contribuyó a crear un clima en el cual se alentaba el periodismo investigativo.

Mientras el aliento al periodismo inquisitivo se remonta al período anterior a Watergate, es indudable que los hechos que se fueron revelando y que culminaron con la renuncia de Nixon, dieron un gran impulso a esta corriente. Watergate fue significativo en este contexto no sólo porque contribuyó a legitimar las actividades del periodismo investigativo (aunque en virtud de la —de algún modo— novelesca actuación de Carl Bernstein y Bob Woodward) sino también porque, por primera vez, llevó al dominio público las regiones ocultas de la cima del poder político donde, sorpresiva e inesperadamente, ellas quedaron a la vista. El contenido que emergió, tanto en términos de la evidencia de actividad criminal y de las impactantemente crudas maniobras de Nixon y sus cómplices, sirvieron sólo para avivar el escepticismo público respecto a la credibilidad y veracidad de sus

líderes. Watergate ayudó a promover un clima de escepticismo en el cual nadie, ni siquiera el Presidente, está libre de sospecha.

El renovado énfasis en el periodismo investigativo, en la década de los años sesenta y setenta, contribuyó a cambiar la cultura periodística y a crear un contexto en el cual la búsqueda de secretos ocultos, y la revelación de ellos cuando eran descubiertos, fue considerada cada vez más como una forma legítima de actividad periodística en los círculos mediáticos. En el contexto de este amplio cambio en la cultura periodística, la distinción entre diferentes tipos de secretos se hizo borrosa y cada vez más difícil de trazar. Una vez aceptado que era posible alzar el telón que resguardaba las regiones superiores del poder, se tornaría muy difícil mantener una distinción nítida entre los secretos relacionados con el ejercicio del poder y aquellos relativos a las conductas de la vida privada. El periodismo inquisitivo caería, así, en una suerte de búsqueda lasciva donde los aspectos ocultos del ejercicio del poder se mezclarían con aspectos ocultos de la vida de los poderosos¹⁵. Los códigos y convenciones periodísticas, que previamente habían alejado a los periodistas de informar sobre la vida y asuntos privados de los líderes políticos, se relajarían gradualmente, y, en algunos casos, se abandonarían del todo, de modo que los periodistas se verían menos limitados por el ethos de su profesión.

Un tercer factor que ha contribuido al aumento de la frecuencia del escándalo político es el cambio en el carácter general de la cultura política. Este cambio tiene sus raíces en una serie de transformaciones estructurales que dieron forma al ambiente donde hoy se realiza la actividad política. Particular importancia ha tenido, en este sentido, la gradual declinación de la política basada en partidos de clase, con sistemas de creencias claramente opuestos y que se enfrentaban en la arena política representando los intereses de las clases que les proveían sus núcleos de apoyo electoral. Durante la segunda mitad del siglo XX, diversos cambios en el mundo del trabajo en las sociedades industriales occidentales, incluyendo el descenso de las industrias tradicionales como las del carbón y acero, y el surgimiento del sector de servicios y una serie de industrias intensivas en conocimiento, han transformado el contexto social de la política. Los partidos ya no pudieron depender de las antiguas clases sociales que alguna vez fueron la base de sustentación. Las divisiones doctrinarias tradicionales se atenuaron y, cada vez más, los partidos y sus líderes tuvieron que luchar por lograr el apoyo de un creciente grupo de votantes no comprometidos.

¹⁵ Véase Michael Schudson, *Watergate in the American Memory: How We Remember, Forget, and Reconstruct the Past* (1992), p. 117 en adelante.

Estas amplias transformaciones sociales nos ayudan a comprender lo que gruesamente podríamos describir como la decadencia gradual de la “política ideológica” y la creciente importancia de la “política de la confianza”. La política tradicional de los partidos basados en clases sociales, con sus sistemas de creencias opuestos y el fuerte contraste entre derecha e izquierda, no ha desaparecido, pero se ha debilitado significativamente debido a las transformaciones sociales del período de posguerra. Y en su lugar ha surgido un tipo de política basada cada vez más en programas políticos específicos ofrecidos por los partidos. El apoyo para estos programas políticos ya no puede obtenerse apelando principalmente a los intereses de clase de los votantes, y la ciudadanía ya no puede contar con que los políticos van a cumplir sus promesas en virtud de las antiguas afiliaciones sociales de sus partidos. Es más, al declinar la antigua política ideológica, muchas personas se sienten cada vez más inseguras respecto a cuál pudiera ser la mejor forma de abordar los complejísimos problemas del mundo moderno; el mundo aparece como un lugar desconcertante en el que no hay soluciones simples y en el que debemos hacer fe en que nuestros líderes políticos serán capaces de juzgar correctamente y proteger nuestros intereses. Es en este contexto donde el asunto de la credibilidad y confiabilidad de los líderes políticos se convierte en un tema cada vez más importante. Las personas se preocupan más del *carácter* y *confiabilidad* de quienes son o pueden llegar a ser sus líderes, porque estos atributos son ahora la principal garantía de que las promesas políticas llegarán a cumplirse y de que frente a la complejidad e incertidumbre se adoptarán decisiones juiciosas. La política de la confianza adquiere importancia no porque los políticos sean hoy inherentemente menos confiables que en el pasado, sino porque las condiciones sociales que antes aseguraban su credibilidad han sido erosionadas.

Este cambio en la cultura política ha contribuido a que los escándalos adquieran un ascendiente mayor en la vida política de nuestros días. Una de las razones por las que el escándalo político se ha vuelto tan gravitante hoy es porque se ha convertido en una *prueba de credibilidad* para la política de la confianza. Cuanto más se orienta nuestra vida política hacia temas que atañen al carácter y la confianza, mayor es la trascendencia que damos a esas ocasiones en las que se cuestiona la confiabilidad de los líderes políticos. Mientras más dependamos de la integridad de los políticos para el cumplimiento de sus promesas y de su capacidad de juzgar con sensatez, más importancia daremos a esas ocasiones en que quedan de manifiesto debilidades de carácter y errores de juicio. Visto de esta manera, podemos comprender por qué un escándalo que concierne a la vida privada

de un político tiene para muchos gran importancia política: no tanto porque ellos crean que los políticos deban adherir a estrictos cánones morales en su vida privada, sino porque están preocupados de qué les dice también esa conducta de su integridad, credibilidad y juicio.

El progresivo ocaso de la política ideológica también conlleva que, en el subcampo de los políticos profesionales, los desacuerdos fundamentales en materia de principios son ahora menos pronunciados y los partidos políticos buscan otros medios para diferenciarse. Así como los partidos se dirigen crecientemente a un centro común y compiten por un conjunto cada vez mayor de votantes no comprometidos, las debilidades de carácter —reales o presuntas— de sus opositores y la infracción de códigos de conducta se convierten en potentes armas en la lucha política. Los asuntos de carácter se politizan cada vez más en la lucha por la diferenciación de los partidos en un contexto donde es más y más difícil apelar a diferencias fundamentales de principio, y donde, en parte para compensar esto, los partidos y sus líderes buscan capitalizar políticamente las debilidades de carácter de los otros.

Puesto que el escándalo pone en tela de juicio la credibilidad exigida por la política de la confianza y puesto que los asuntos de carácter se encuentran más ligados a la lucha partidaria, la irrupción del escándalo tiende a tener un efecto acumulativo: el escándalo alimenta al escándalo, precisamente porque cada escándalo hace que la atención se centre aún más en la credibilidad y confiabilidad de los líderes políticos. El efecto acumulativo del escándalo se incorpora en la carrera electoral en la medida en que los partidos políticos y sus líderes potenciales construyen sus campañas políticas en base a la denuncia de episodios pasados que ponen en tela de juicio la credibilidad de sus contrincantes. De ahí que en el período posterior a Watergate, Jimmy Carter construyera su campaña presidencial de 1976 sobre el tema de la confianza y prometiera que su administración restauraría los altos estándares éticos del gobierno. “Nunca les mentiré” fue su famosa declaración de campaña. Sin embargo, tan pronto como asumió el cargo, una ola de nuevos escándalos salió a la luz. Carter llegó a la Casa Blanca con la promesa de realizar un gobierno honesto, decente y limpio, pero dentro de una cultura política centrada en el carácter y la confianza, esto fue como agitar una bandera roja frente a las huestes de críticos y periodistas.

Bill Clinton se vio atrapado en un círculo similar, aunque en forma mucho más acentuada. Como tantos aspirantes a la presidencia en el pasado, Clinton hizo su campaña prometiendo eliminar de la política las prácticas indeseables del gobierno de Reagan, y realizar “la administración más

ética de todas". Pero pronto se encontró con que miembros de su propio gobierno, y, por cierto, él mismo y su esposa, estaban siendo investigados por posibles conductas financieras inapropiadas. También se encontró con que acusaciones y revelaciones concernientes a su vida privada se convertían en asuntos públicos, lo que casi malogró su campaña de 1992 (por su relación con Jennifer Flowers) y terminó en un proceso y juicio de destitución por el Senado a raíz de la revelación de su relación con Mónica Lewinsky. Por supuesto que lo que llevó al proceso de destitución de Clinton no fue la revelación de este hecho en sí, sino una serie de otras transgresiones de segundo orden cometidas en relación con un caso de acoso sexual alegado por Paula Jones, en cuyo contexto Clinton dio testimonio bajo juramento negando haber tenido relaciones sexuales con Mónica Lewinsky, y por consiguiente exponiéndose al cargo de perjurio, entre otros. Pero lo que emergió en el curso de esta larga saga fue la revelación, sin precedentes, de las relaciones privadas entre el Presidente y una joven en práctica al interior de la Casa Blanca, una revelación que impactó y horrorizó a mucha gente (a pesar de que muchos creyeron, al mismo tiempo, que el escándalo había sido llevado fuera de toda proporción por unos medios de comunicación agresivos, un equipo investigativo insistente y un Congreso parcial).

He usado el ejemplo del escándalo político para ilustrar la transformación de la naturaleza de la visibilidad, tratada a propósito del desarrollo de los medios de comunicación. Vivimos hoy en una era de alta visibilidad mediática, y aquellos que detentan o aspiran a posiciones de importancia en la vida pública se enfrentan a un ambiente informativo muy diferente de aquel que existía algunos siglos (e incluso algunas décadas) atrás. Debido al desarrollo de múltiples formas de comunicación mediática y al surgimiento de numerosas organizaciones mediáticas relativamente independientes del poder estatal, los actores políticos deben actuar hoy en un ambiente informativo más intenso, más extenso y menos controlable del que había en el pasado. Es *más intenso* en el sentido de que la cantidad total de flujo informativo es mucho mayor que antes, así como más y más organizaciones y redes comunicacionales hacen disponible un siempre creciente volumen de material simbólico. Es *más extenso*, en el sentido de que la variedad de individuos que son alcanzados por estas redes de comunicación y capaces de recibir el producto de las organizaciones mediáticas, es mucho mayor de lo que era un siglo (o incluso algunas décadas) atrás. Así la distribución geográfica de estos receptores es mucho mayor: hoy la información fluye muy rápido a través de redes que son no sólo nacionales sino crecientemente globales en su alcance. Y el ambiente informativo es *menos controlable* en el sentido que, dada la proliferación de formas mediáticas de comunica-

ción, es mucho más difícil para los actores políticos tejer un velo de confidencialidad alrededor de sus actividades y predecir las consecuencias de revelaciones no deseadas. Quiéranlo o no, hoy los líderes políticos pueden ser vistos por un mayor número de personas y escrutados más de cerca que nunca antes en el pasado y al mismo tiempo están más expuestos al riesgo de que los actos que realizan o las palabras que expresan en forma encubierta o privada sean reveladas al público.

REFERENCIAS BIBLIOGRÁFICAS

- Bobbio, Norberto. *Democracy and Dictatorship: The Nature and Limits of State Power*. Traducido al inglés por Peter Kennealy. Cambridge: Polity, 1989.
- Burke, Peter. *The Fabrication of Louis XIV*. New Haven, Conn, y Londres: Yale University Press, 1992.
- Darnton, Robert. *The Forbidden Best-Sellers of Pre-Revolutionary France*. Londres: Harper-Collins, 1996.
- Elliott, J. H. "Power and Propaganda in the Spain of Phillip IV". En Sean Wilentz (ed.), *Rites of Power: Symbolism, Ritual, and Politics since the Middle Ages*. Philadelphia: University of Pennsylvania Press, 1985.
- Geertz, Clifford. "Centers, Kings, and Carisma: Reflections on the Symbolics of Power". En su obra *Local Knowledge: Further Essays in Interpretive Anthropology*. Nueva York: Basic Books, 1983.
- Habermas, Jürgen. *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society*. Traducido al inglés por Thomas Burger con la asistencia de Frederick Lawrence. Cambridge: Polity, 1989.
- Hagood, Wesley O. *Presidential Sex: From the Founding Fathers to Bill Clinton*. Secacus, NJ.: Carol Publishing Group, 1998.
- Jamieson, Kathleen Hall. *Eloquence in an Electronic Age: The Transformation of Political Speechmaking*. New York y Oxford: Oxford University Press, 1988.
- Mercurius Britannicus*, 92, 4 de agosto de 1645, p. 825.
- Meyrowitz, Joshua. *No Sense of Place: The Impact of Electronic Media on Social Behavior*. Nueva York y Oxford: Oxford University Press, 1985.
- North London Press*, 16 de noviembre de 1989, p. 5.
- Schudson, Michael. *Discovering the News: A Social History of American Newspapers*. Nueva York: Basic Books, 1978.
- Schudson, Michael. *Watergate in the American Memory: How We Remember, Forget, and Reconstruct the Past*. Nueva York: Basic Books, 1992.
- Slevin, James. *The Internet and Society*. Cambridge: Polity, 2000.
- Thompson, John B. *The Media and Modernity: A Social Theory of the Media*. Cambridge: Polity, 1995. [*La Media y la Modernidad: Una Teoría de los Medios de Comunicación*. Barcelona: Paidós, 1998.]
- Thompson, John B. *Political Scandals: Power and Visibility in the Media Age*. Cambridge: Polity, 2000. [*El Escándalo Político: Poder y Visibilidad en la Era de los Medios de Comunicación*. Barcelona: Paidós, 2001.]
- Whitaker, Reg. *The End of Privacy: How Total Surveillance is Becoming a Reality*. Nueva York: The New Press, 1999. □

DEL HUMOR, DEL DOLOR, Y DE LA RISA (CRÓNICA DE UNA DEPRESIÓN)

Alfredo Bryce Echenique

...Como el humorista triste que sale al sol de la convalecencia y la salud recuperadas, a reírse un poco y bastante de su tremenda depresión, así también el desesposado de esta crónica salió a reírse abundante y exageradamente de su larga, muy larga depresión neurótica. Recogió todo lo que de ella pudo recoger e inventó con su memoria todo lo que de ella pudo inventar, como homenaje a la vida dolida, a un médico maravilloso llamado Z., a un amor perdido...

Nunca sabré cómo llegó a mis manos el frasco de Tranquilizante-1000 que llevaba conmigo en aquel viaje a la Costa Azul. Tampoco sabré nunca por qué llegó a unas manos —las mías— nada habituadas al consumo de pastillas de ningún tipo, y que más bien las rechazaba, debido al uso y abuso que mi madre había hecho de ellas —sobre todo de las pastillas para los nervios, como se les llamaba en casa—. Para mí, las grandes

ALFREDO BRYCE ECHENIQUE. Escritor. Licenciado en Derecho y Doctor en Letras por la Universidad de San Marcos de Lima. Ha enseñado en distintas universidades; entre ellas, la Universidad de Nanterre, la Sorbona, Vincennes y la Universidad de Paul Valéry de Montpellier. Autor de las novelas *El Huerto de mi Amada* (2002), *Un Mundo para Julius*, *Tantas veces Pedro*, *La Vida Exagerada de Martin Romaña*, *No me Esperen en Abril* y *La Amigdalitis de Tarzán*, entre otras, y de las colecciones de cuentos *Magdalena Peruana* y *Guía Triste de París*.

situaciones de angustia eran producto de una fuerte timidez, y sobre todo de una timidez que no se manifestaba como parquedad al hablar, sino más bien como todo lo contrario. Yo llenaba el universo entero de palabras, anécdotas, y de muy inverosímiles historias, para ocultarles a sus habitantes, uno por uno, el feroz temblor de mis manos, los tremendos calambres y los atroces dolores de cabeza que determinadas situaciones me producían. No era, pues, un tímido a tiempo completo, por decirlo de alguna manera, sino un hombre que, de ser enteramente dueño de una situación, de pronto era atacado por una suerte de inesperado trastorno que lo llenaba de pánico y lo dejaba reducido a la condición de mísero tembleque al que todo se le caía de las manos, pero que, al mismo tiempo, era capaz de narrar, de principio a fin y sin alterarles una sola palabra, *Las mil y una noches*, en un loco y desesperado afán de no ser descubierto temblando y aterrado. Y a nadie he odiado tanto desde entonces como a esas personas que resuelven el expeditivo timidez con un cómodo y profundo silencio que, además, generalmente viene acompañado de una mirada lánguida al personal y de un par de ojos bien grandes y generalmente oscuros, que el mundo —el mismo detestable e infecto mundo que a mí me aterra y reduce a temblor crónico— suele encontrar hermosos, penetrantes, profundos, inteligentísimos, enamorados del bien y de mí —y/o espejos de un alma que esconde tesoros secretos y secretos tesoros—. En fin, todo esto mientras mi mal y yo no tenemos remedio alguno, según me hizo saber nada menos que un brillante psicólogo peruano que me estaba sometiendo a un brillante interrogatorio internacional para optar a una nada brillante beca francesa. Satisfecho por mis respuestas a sus preguntas, el doctor L. acababa de extraer un cigarrillo nacional y se disponía a encenderlo, cuando yo, también satisfecho por mis respuestas a sus preguntas, aunque creo que al revés, en mi caso, o sea más bien satisfecho por las preguntas que facilitaron mis respuestas y la beca a Francia, extraje también un fósforo, y lo encendí, pero la llama ardiente arrancó a temblar en el aire, entre la corbata del doctor L., su prominente nariz aún joven, o sea que nada eminente aún, como fácilmente podemos deducir, y el maldito cigarrillo que, oh espanto, temblaba casi al mismo ritmo que mi mano le estaba imponiendo e incluso infundiéndolo al mundo, pues la situación empezaba a ser aterradora, pero mi mano dale con temblar y el fósforo dale con apagarse y el doctor L. dale con seguir sin lograr fumar. Y así de pavoroso, todo, hasta que aquel grande y buen narigón me dijo:

—Yo también temblaba así a su edad, joven. Y míreme usted, ahora.

¡Qué bárbaro! ¡Cómo temblaba el doctor L. y qué inteligente y qué generoso era de explicármelo todo así, tan claramente y para siempre! Y,

además, mientras literalmente me decía que me fuera con mi propio temblor a Francia.

Y en Francia andaba, ahora, yo, y en la Costa Azul, y en un hotelito de Saint-Raphaël, muy precisamente, cuando la esposa me dijo: “Apaga ya la luz, y déjate de cuentos”. Le dije que no me atrevía a apagar la luz y a dejarme de cuentos, y ella me respondió apagando por mí y explicándome que mañana me iba a acompañar, sí, bueno, qué pesado eres, pero sólo hasta la estación del tren para que yo luego siguiera mi viaje en solitario hasta Barcelona, donde me esperaba la luz de una esperanza. La esposa se quedó dormida en pleno nuestro primer y quizás último viaje a la Costa Azul, qué despilfarro, y mi angustia y mi recién estrenado terror a la oscuridad empezaron a crecer sin explicación alguna. Y sin siquiera un antecedente alguno, por la sencilla razón de que yo nunca antes había sentido miedo por causa de la oscuridad. Ni tampoco por causa de la claridad ni de la lluvia ni del sol ni de la primavera, el invierno, el verano o el otoño. Estaba estrenando ansiedades y terrores y la esposa sencillamente no lo podía creer, y, como tantos con tantos otros inexplicables y muy recientes estrenos, se indignaba al verme así, a menudo indigno, indigno y sin respuestas. Sólo aterrado por algo que nunca antes lo había aterrado a uno. Aterrado en un cine, aterrado al ir a mis clases, aterrado de estar aterrado. Y con esos ataques de terror en medio del terror, para colmo de males, que debe de haber sido entonces cuando un alma caritativa me entregó aquel frasco de Tranquilizante-1000 que yo solía abrir y consumir en un abrir y cerrar de ojos, como quien le da a su atroz alarido el alimento debido. Pobrecita, la esposa, qué indiferente dormía a mi lado, qué inocente de mi dolor, qué de espaldas al mundo de todo dolor que no fuera de muelas y asociados, digamos. Y cómo, enamorado y respetuoso, se aguantaba uno el alarido o lo taponeaba con una mano feroz mientras encontraba la puerta de una habitación que generalmente daba al vacío y se escapaba corriendo con el frasquito color serenidad apretado con toda el alma por una mano que, ya más de una vez, en vez de tener cinco dedos, tuvo siete, mientras que uno, siempre con ese sentido del humor tan de uno, se dijo, Claro, pero si es un trocito de la otra mano que viene a ayudarme a abrir el frasquito. Mas resulta que la otra mano acababa de desarrollar un sentido del humor realmente colosal, porque, en vez de los dos dedos que hacían el papel de extras en la historia, se aparecía nada menos que con los diez dedos de las dos manos y por lo tanto había siete dedos de más y la esposa tan dormidita y de espaldas y gracias a Dios, eso sí, que uno ya había alimentado bien a su alarido hasta la próxima vez, que sería cuando, y no Sabe Dios cuándo sería, que es algo muy distinto a Sería cuando, de la misma manera en que

entre este alarido taponeado y el próximo las intensidades se habrían duplicado, y por intensidades quiero decir el pavor, la ansiedad, la bestial angustia y su fatal compañero el temblor que va por dentro y se instala cada vez más cómodamente en nuestra muerte en vida.

Porque la depresión ha llegado sin antecedentes y sin cultura alguna sobre ella, sin experiencia que valga, y sin mecanismo alguno de defensa contra ella, o tal vez sólo ese humor que nos permite observarnos indignos de la esposa, mientras ella cree que nos hemos empequeñecido, achicado, acobardado, ante la despedida y ante el mundo en general en la estación de Saint-Raphaël, esa mañana, cuando sólo el ingrediente humor —oculto y como muerto, por ahora, pero también agazapado como un paparazzi detrás de ese árbol—, saca mil placas de la escena en la que no sólo me he rebajado para inspirar cariño, el último en esta miserable vida en común adorándonos, sino también para observar mejor cada detalle de nuestra decadencia y caída y cada detalle también de la *Nada* interpuesta con su color blanco como la ballena aterrador de Herman Melville, sí, para observarlos mejor y ayudarnos en el cuento del día en que seremos capaces de recuperar nuestra real estatura y con ella nuestra dignidad. Y, en efecto, en la primera fotografía que ese paparazzi nos mostró, tiempo después, de aquella despedida en Saint-Raphaël, era la esposa la que lloraba y lloraba, mientras me abandonaba, y éramos yo, el enfermo, y su frasquito de Tranquilizante-1000, quienes la consolaban mientras *ella* nos abandonaba. ¿Resulta cuando menos divertido, no? ¿Y resulta paradójico, no...? Por algo, claro, el escritor argentino Julio Cortázar, que tan a menudo había recurrido a situaciones *así* de inexplicables, se había referido más de una vez al “lado cómicamente grave de la realidad.”

Bueno, sí, ¿y por qué no...? Por qué no si ya decía Miguel Mihura que “El humor es verle la trampa a todo”. Y, por su parte, Georg P. Burns afirmaba que: “Quien nos hace reír es un cómico. Quien nos hace pensar y luego reír es un humorista”. Y, por su parte, Miguel de Cervantes escribía: “No son burlas las que duelen, ni hay pasatiempo que valga si son con daño de tercero”. Y escribía también Ramón Gómez de la Serna: “El humor hace pariente de la mentira a la verdad, y a la verdad de la mentira”. O sea que: “Si vais a la felicidad, llevad sombrilla”. Y ya hay hasta quien se ha atrevido a escribir, como el gran Antonio Fraguas, alias “Forges”, que “Si todos los estados reconocieran el derecho a la sonrisa como uno de los derechos humanos, se vería que tal vez al existir este derecho a la sonrisa todos los demás estarían incluidos”.

Pero sigamos con esta crónica de una depresión vista con humor por todas partes, para que duela menos revivirla pasito a paso. La esposa está

llegando en un tren a París, mientras que él, a quien desde este momento llamaremos el desesposado, está tapando con las justas su frascuito de Tranquilizante-1000, elevado sobre un andén de la estación de trenes de Barcelona, y sólo cuando logra tapanlo bien logra asimismo poner los pies sobre la tierra, por fin, y saca de un bolsillo una agenda, mientras, *prácticamente al mismo tiempo*, y dándole a este detalle una Gran Importancia, *La Mayor Importancia*, comprueba que los cinco dedos de cada una de sus dos manos suman diez, pero de pronto mira y se desangra, porque sí: son once. Y se consuela al recordar que bueno, que en Saint-Raphaël llegaron a ser multitud y que hasta había uno con pelo rojo, y se sonríe en medio de su profunda enfermedad del alma porque ya va aprendiendo que, así como en el sencillo y envidiado y común resfriado, hay altos y bajos, y en los altos trenes yo voy con ella y en los bajos desciendo yo sin ti, sí, en la vida se aprende siempre, y puede ser bendita y malditamente, como en el humor y en la ironía, y así, aprendiendo un poquito más y sabiendo que sus once dedos pueden ser otra vez multitud, y entre ésta uno pelirrojo, encuentra por fin el teléfono del médico de mi alma —él aún no se da cuenta del acierto, de que se les llama médicos del alma a los psiquiatras—, y sí, sí tiene cita, y es esta tarde, sí, señor, y él agradece en el alma que se hayan acordado de él, porque el ahora también desesposado ha logrado llegar solo a una ciudad y no morirse y marcar un número en el teléfono y que le contesten, y preguntar por una persona y que le digan Sí, dar su nombre y que le digan que es esperado a tal hora, en fin, y todo justito ahora en que el frascuito de Tranquilizante-1000 se me quedó vacío. Entonces el desesposado siente que aferrarse a un frascuito vacío en una ciudad que no conoce, pero donde es esperado a una hora en punto, le devuelve cinco dedos por un buen rato a cada una de sus manos y esto lo llena de esa paz interior y triste que luego hay que expulsar agradecida y tiernamente por los rincones.

Y así resulta que ha llorado tanto para cuando llega la hora de la consulta, el desesposado, que ahí lo tenemos ahora, sentado en una habitación severa, de muebles igualmente severos, pero ante un médico que él francamente encuentra como muy poco apropiado para la situación. Porque el doctor Z., que así se llama este médico, viste juvenil y alegremente, y acaba de recibirlo con un jovial apretón de manos. En fin, cuestión de estilo, probablemente, pero uno siempre preferiría que lo trataran de acuerdo a su condición, o sea con la debida gravedad, en este caso. Y es que el desesposado acaba de soltar, cubriéndose de ridículo ante su propia memoria del presente y del futuro, que: Aunque a usted no le parezca, doctor Z., mi mal no tiene remedio.

Pero el doctor Z., que lleva un buen rato observándolo con una sonrisa no desprovista ni de humor ni de ternura, le dice a boca de jarro que lo que va a hacer, para empezar, es responderle a una pregunta que usted, señor desesposado, está desesperado por hacerme, pero no se atreve. Y el desesposado se rasca y rasca la cabeza pero no, nada de salirle la pregunta esa, que ahora sí, por fin, sale de labios de un doctor que, él lo sabe, acaba de salvarle la vida por primera vez. Pero ¿cuál es la dichosa pregunta, doctor? Pues que usted no se va a suicidar, amigo. La palabra *amigo*, escuchada así, en una ciudad que no conoce, de boca de un hombre al que acaba de conocer, sin su frasquito de Tranquilizante-1000, y con su lejana y adorada esposa allá, rumbo a París, cada vez más lejana, cada vez abandonándolo más, o acompañándolo menos, en todo caso, la palabra *amigo* también le ha servido al desesposado como una boya, y como mil boyas, por ser una característica de la depresión el ir de amigo en amigo como un náufrago va de boya en boya, aunque sin dejarlo saber jamás, sin agradecerlo siquiera, sólo náufragamente, que es así como vive y va por el mundo quien sufre una buena depresión.

Bien, pero hay también los terrores inesperados, la oreja gigantesca de un señor que va por la calle sin que jamás nadie se fije particularmente en él, sólo uno, sí, sólo uno, y a centenares de metros de distancia. Como hay los dedos de la mano que sobran y además se nos aparecen pelirrojos, o las idas al cine que debemos interrumpir porque no bien se apaga la luz la angustia a borbotones empieza a estrangularnos en nuestra butaca. Por eso siempre buscamos un asiento de pasillo, y lo más cerca posible de cualquier puerta que diga EXIT o ESCAPE, sí, y es que a cada rato nos vamos a tener que escapar perseguidos por nuestros pavores, por esos pavores que no vienen precedidos por el más mínimo aviso, pero que nos dejan empapados y deshechos por los rincones. Sí, por lo rincones, y sí, porque los rincones son como trampas de inconmensurable imantación que la vida nos va colocando hasta en las rectas más interminables. Y sí, porque el hombre deprimido camina a trompicones por ciudades que siempre le serán extrañas, sumamente extrañas. Tanto que la propia ciudad en que nacimos y crecimos puede resultarnos extraña a los seres que hemos sido invadidos por *El mal oscuro*, como llamó a la depresión el gran novelista italiano Giuseppe Berto, y también a los seres que somos presa de *Los fantasmas de mi cerebro*, como a su vez llamó a la depresión el novelista español José María Gironella. Y a la depresión, creo yo, se refirió también el gran bardo popular que fue el argentino Atahualpa Yupanqui, cuando habló de que a *uno lo moja hasta 'dentro, la oscuridad...*

Pero yo me estoy apoyando en una depresión personal, o sea siempre en un hecho cuyo recuento prefiero hacer desde la propia experiencia, y apoyándome a veces en un anecdotario que no por carecer de una apropiada terminología científica, por ejemplo, deje de tener su relativa importancia y, al menos para mí, una total veracidad. Por ejemplo, es muy cierto que para mí fue indispensable la confianza absoluta que, desde el primer momento, se estableció en el trato y en la relación de varias décadas (hasta su fallecimiento) entre el doctor Z. y yo. Podía ser la confianza de dos pecadores, y de dos hombres de muy distintas formaciones y orígenes y nacionalidades, pero era, digo y repetiré, confianza antes que nada y a prueba de balas. Y existió también una importante diferencia de edades, pero creo que a medida que yo pueda seguir enumerando las diferentes distancias que hubo siempre entre ese gran médico y yo, lo único que continúa creciendo más y más, incluso ahora que escribo y que hace años que él está muerto, lo único que crece y crecerá siempre es esa total confianza basada en un afecto que tuvo indudablemente de afecto paternal, filial, y de afecto fraternal. Y hasta me atrevería a decir que entre el gran médico que fue mi amigo Z. y yo, también tuvo una profunda importancia en la terapia (al menos así lo viví yo) una suerte de profunda y multilateral complicidad. Y completo esta idea agregando que, en todo caso a mí, me ayudó muchísimo a salir de aquella depresión feroz que él calificó de neurótica, el hecho palpable y muy real de que, desde el primer momento, o casi, nos presentáramos el uno ante el otro con todas nuestras cartas sobre la mesa y, diría yo, desnudos, sí, con el alma desnuda.

Todo esto fue importantísimo para un muchacho (yo entonces andaba por los 27 años) que *estrenaba y se estrenaba* en Europa, que venía de un hogar lleno de dolor callado, siempre, de silencios eternos y de tabúes que se sacrificaban ante el altar de la honorabilidad, del buen ser y el mejor parecer, en una ciudad que ya mostraba los mil muñones de la miseria y que ya formaba parte también de esa “América descalza, que habla el español de pedir y mendigar”, según las duras palabras del gran escritor puertorriqueño Luis Rafael Sánchez. Yo era una suerte de heredero del silencio y la buena salud, de las buenas conciencias y los mejores modales y educación. Y, aunque esto se lo debo a mis padres y abuelos resulta que yo no salí así, o sea que resulta que yo salí así, o sea al revés y culposo y, ¡oh gran colmo de males!, demasiado sensible como para tragarme tanto machismo. Y con coraje y hombría y navegando a contra corriente me hice escritor pero resulta que no me hice banquero pero resulta que no por eso dejé de adorar a mis padres y abuelos y hasta a agradecerles por ese bien suyo que ellos habían querido poner en mis manos y en mi futuro *per vitam aeternam*.

El tremendo culposo enclenque y tembleque que recibió el doctor Z. en su despacho de Barcelona —y qué bien lo vio él— sólo reaccionaría a los estímulos del afecto, del gran cariño, incluso, y de toda esa monumental confianza que había sin duda que otorgarle para ir paliando poco a poco los atroces efectos de una soledad abrumadora, de la culpa inmensa de sentirse *otro*, de una lucidez a veces arrolladora, pero las más de la veces empapada de culpa y tremendamente generadora de desasosiego y más culpa aún, de la demencia y su dolor-locura, de una gigantesca ambición literaria pero ninguna vanidad para ponerla en práctica mucho más allá de los límites reales de un vicio oculto, y una capacidad de entrega a la juerga de la vida sólo comparable al deseo paralelo de amar y ser amado y cuyos límites, gracias a Dios, existían y se ejercían muy a menudo en la permanente necesidad de cumplir con su deber, en una gigantesca capacidad de orden que apagara las llamaradas de los grandes abrazos, en una maniática puntualidad y en una minuciosidad para los detalles y pequeñas cosas y objetos de la vida cotidiana que en mucho semejaban a las menudas tareas que han servido tanto de terapia a las personas que viven un internamiento real, o sea de manicomio.

Yo vivo con la convicción profunda de que el doctor Z. vio en mí todas estas realidades y tendencias o posibilidades desde la tarde noche de nuestra primera cita en Barcelona. Y a ello adjuntó —digamos— unos míticos folios que me pidió redactar en mi cuarto de pensión, esa misma noche, de tal manera que se los pudiera entregar en manos propias o hacérselos llegar en la mañana del día siguiente, un sábado, para llevárselos con él de fin de semana. Apenas volvimos a hacer mención de aquel puñado de folios que yo escribí como loco y del cual no he logrado nunca recordar una línea. Él decía que era una novela. Y él sabía mucho más que yo, pues era un gran lector y yo no era más que un debutante enfermo con sus culpas y contrastes, con su enorme pena de amor, pero también con su sentido innato del humor. Lo más probable es que en ese puñado de papeles estuviese el embrión de una primera novela. Aunque no es esto lo que me importa. Y ni siquiera es lo que él me dijo y yo a veces he creído. En realidad estoy seguro de que aquellos papeles eran un inmenso pedido de confianza y de afecto. Y a este pedido, estoy seguro, se unía otro que consistía en reclamarle al doctor un espacio para la risa y el humor —no para la comicidad ni la carcajada— en toda nuestra relación de médico y paciente y en toda aquella otra relación que debutaba esa tarde-noche y que —creo—, él y yo presentimos que iba a durar para siempre.

Y lo demás fue la mencionada confianza, la distancia obligatoria de aquel tratamiento en que mucho se hizo a través de cartas que iban desde su

consulta de Barcelona hasta mi departamento de París y que sólo se interrumpían cada vez que había vacaciones universitarias, ya que por aquellos años yo trabajaba dictando clases de literatura y civilización latinoamericanas en la Universidad de Nanterre, y sólo me era posible ver al doctor Z. cuando las clases se interrumpían. En estos casos, yo tomaba siempre el primer tren que saliera en dirección a la Ciudad Condal, y, no bien llegaba, llamaba siempre al consultorio y pedía mi cita como cualquier otro paciente, aunque siempre a sabiendas de que, terminada la consulta, el doctor Z., su esposa e hijos, me invitarían a cenar esa misma noche. Y, aunque sin entrar en los detalles más o menos profesionales, en cierta medida las conversaciones de sobremesa en numerosos restaurantes de Barcelona eran una suerte de comedia y más discreta continuación de algunos problemas que habíamos abordado poco antes en el consultorio. Y fue de esta manera cómo el doctor Z. y su familia se convirtieron para mí en la encarnación de una propia familia, pero en la cual había ingresado yo siendo ya mayor de edad. Y hoy que él ha fallecido, yo continúo viendo todo lo que puedo a su viuda e hijos. En fin, he insistido mucho en este aspecto amical de la relación médico-paciente entre aquel gran psiquiatra y formidable hombre, porque estoy convencido de que, al menos en mi caso, fue realmente indispensable. Y no sólo porque en aquellos años tan duros para mí necesitara de ese cariño, de esa familiaridad y de esa confianza, sino porque además realmente me equipó y me dio una serie de elementos que me han sido indispensables para enfrentarme a otros *bajones*, término éste tan frecuente entre las personas que sabemos hasta qué punto la depresión se aleja y hasta parece olvidarnos por completo, durante largas temporadas, pero finalmente no nos suelta del todo jamás.

Pero, antes de volver a la depresión *per se*, quiero despedirme de mi gran amigo e inolvidable médico y persona, el doctor Z., contando una anécdota muy divertida que, además, creo que lo retrata bastante bien, y, en todo caso, dice mucho de la gran confianza que hubo siempre entre nosotros. El doctor Z., consciente de que yo entonces vivía —a todo nivel— un largo período de vacas flacas, solía llenarme los bolsillos de todas las pastillas posibles que yo pudiese necesitar, allá, solo en París. Me atiborraba los bolsillos, la verdad, pero resulta que un día él tuvo un colerón telefónico de padre y señor mío con su cuñada, de nacionalidad suiza. Ésta lo había acusado muy torpemente de haber fracasado por completo con la depresión de su hermano muy querido, y el doctor Z. había pescado tal colerón que hasta miedo nos daba por su salud a su esposa y a mí. Y hubo que arrancarle el auricular de entre sus rabiosas manos y colgarle a la cuñada suiza, cuando a gritos la estaba acusando de ser ella, por el contra-

rio, suiza de eme, la que le estaba matando a su hermano a fuerza de ensaladas de lechuga, en vez de dejarlo que se coma una buena paella regularmente, como un español cualquiera, en fin, que hubo que cortar por lo sano aquel colerón, cuando el pobre doctor Z. se dio cuenta de que necesitaba un Tranquilizante-1000, y hasta dos, y fui yo quien tuvo que administrárselas, de entre las mil cajitas y frasquitos con que acababa de llenarme los bolsillos. Y, perdonen la apostilla, pero un médico así me curará siempre a mí, por el hecho único de ser una persona, muy humana, maravillosamente humana.

Y ahora, volviendo a la enfermedad, así como en el caso que narro el desesposado cansaba a la esposa, que había dejado precisamente de ser *su* amante esposa, para convertirse sólo en *la* esposa, siendo ellos una pareja de recién casados, y sólo por amor, y por un gran amor, además, así también creo yo que a los seres que hemos vivido *la* depresión y vivimos siempre bajo *su* amenaza, bajo su espada de Damocles, nos acompaña un cierto grado de incompreensión real o potencial, entre los seres con que vivimos y a quienes amamos. Y ello nos hace a menudo ocultarle más de un *bajón* a más de una persona de nuestro entorno más cercano. Por ello me encanta la frase aquella del francés longevo y genial que fue Fontenelle, cuando un médico que lo visitó en plena “postración”, le preguntó qué sentía. A lo que el sabio respondió: “Lo que siento, doctor, es una cierta dificultad de ser. Nada más.”

¿No es ésta, acaso, una de las mejores definiciones de lo que es, por lo menos en sus comienzos, una depresión? Y luego, cuando la enfermedad se agrava, cuando *el mal oscuro* se agrava, ¿no se convierte esa cierta dificultad en una dificultad cierta y total? A mi modo de ver, la depresión es principalmente una auto-agresión, una feroz autocrítica y permanente y cruel auto-observación que nos paraliza hasta tumbarnos totalmente. Se dice que los depresivos se convierten en seres que se ocultan incluso debajo de sus sábanas, pero a mí me consta que este mal puede agravarse hasta convertirse el enfermo en un mudo aterrado que se oculta debajo de su cama. Y conocí el extremo, sí señores, de un poeta elegante y sensual, amante de las formas, e incluso un ser para el cual el absurdo de la vida se salvaba exclusivamente por las formas, gracias a éstas. Diré más, aún. Diré que la depresión puede llegar a extremos en que se convierte en un asunto ligado a nuestro presupuesto.

Pues sí. Me consta. Yo visitaba a aquel querido amigo y gran poeta en su ataúd de enfermo. No en su lecho de enfermo sino en el ataúd elegantísimo que se había mandado hacer para *morir* o *postrar* ahí adentro sus muy largos períodos de depresión. Sus amigos nos sentábamos horas a

visitar a aquel muerto en vida, que, lo juro, jamás pronunció una palabra ni miró a sus amigos, porque sabe Dios adónde miraba, o si no es que estaba mirando directamente el fondo de la noche del dolor, autoagrediendo ahí, tan elegante como desesperado. Y recuerdo también que, durante una recaída, el desesposado de esta crónica se pasó una larga temporada de perro metido debajo de su cama. Y a veces pienso que lo hacía incluso por una asunto presupuestal. ¿Que cómo? Pues, digamos, que así. Miren. El desesposado, hombre amante de las formas elegantes, pero de recursos más bien muy escasos, sobre todo por entonces, jamás hubiera podido costearse un ataúd elegantísimo y casero para períodos de postración. Pero, en cambio, era, como tantos otros depresivos, hombre de humor y grandes ironías cervantinas. Quiero decir con esto que no se burlaba del dolor ni de los defectos ajenos sino de las virtudes personales o ajenas, como ocurrió con Cervantes y don Quijote. Y como el desesposado, además de hombre de humor era amante de los perros, había escogido la parte de debajo de su cama para los períodos en que su vida era una vida de perros. Y así, de esta manera, en su dormitorio de persianas siempre cerradas, en su cámara negra, siempre había un momento en que miraba hacia un lado y comprobaba que, como a los perros que amaba, la esposa le dejaba diariamente un platillo de comida ahí en el suelo, junto a la puerta que apenas entreabría ella, de ese velorio en vida.

Tengo para mí, y “cosas” he leído por ahí, que la depresión, cual enfermedad literaria y realmente vengativa, visita sobre todo a los humoristas¹. Y puedo contar, por ejemplo, que una vez en Barcelona estuve presente en una improvisada reunión de los más grandes humoristas gráficos de la España de entonces, de hoy, y de siempre. Sí, señores, aquellos mismos humoristas cuyos nombres o firmas probablemente se les están viniendo a ustedes a la cabeza en este mismo momento. Para empezar, esos hombres profundos y tan serios habían escogido, para reunirse, casi el sótano oscuro y desierto de un triste bar, más que un piso bajo o una historiada cava. Nunca se rieron. Nunca nos reímos. Nunca, mientras tuve el honor de estar con ellos, de ser presentado, con muchísimo honor y orgullo y placer de mi parte, me reí. Y, sin embargo, con la distancia de los años, de unos buenos treinta años, diré que aquella fue una noche llena, de la misma manera en que, una vez al mes, la luna es luna llena.

Y aquí vuelvo a recordar al gran Julio Cortázar, irreverente, cronopio inmenso, humorista genial que en aquellos setentas y ochentas se queja-

¹ Véase el excelente ejemplar que, con motivo de su trigésimo aniversario, le dedicó al humor y a la risa *Jano*, la revista mensual de “Medicina y humanidades”, número extra de octubre de 2001, Vol. LXI, N° 1.406. De la página 108 de este mismo ejemplar, he sacado todas las citas sobre el humor que aparecen en este texto.

ba del exceso de gravedad de la literatura latinoamericana, y de que, como decía él, “En América latina, no bien un hombre empieza a escribir, se pone serio”. Y agregaba, el autor de mil geniales cuentos literariamente subversivos: “¿Hasta cuándo será el humor patrimonio exclusivo de los anglosajones y de Jorge Luis Borges?” Pues hasta él mismo, por supuesto, hasta el mismo gran Cortázar, mi maestro, que, como dije antes, se ocupó ante todo del “lado cómicamente grave de la realidad”. Y en los libros del gran Cortázar, como en los de tantos otros grandes cultivadores del humor, neurosis y depresión se leen muy a menudo entre líneas y en las propias líneas. También Anatole France escribía: “Es posible que me hubiera aniquilado la tristeza, si no me reanimase la facilidad para descubrir la parte cómica de las cosas”.

Y, así también, como Giuseppe Berto en su ya citada novela *El mal oscuro*, el humorista triste sale al sol de la convalecencia y la salud recuperadas, a reírse un poco y bastante de su tremenda depresión. Y así, también, tantos y tantísimos humoristas más. El propio desesposado de esta crónica salió, asimismo, alguna vez, en algún libro de los suyos, a reírse abundante y exageradamente de su larga, muy larga depresión neurótica. Recogió todo lo que de ella pudo recoger e inventó con su memoria todo lo que de ella pudo inventar, como homenaje a la vida dolida, a un médico maravilloso llamado Z., a un amor perdido, a la infinita pequeñez de la que fue capaz, y así, de esta manera, logró recuperar, al menos en la literatura suya, que intenta siempre digresivamente incorporar al libro el caos y desorden de la vida, sus grandezas y miserias desmesuradas; y logró recuperar, repito, la estatura moral perdida y también el amor por unos años de su propia vida que creía perdidos para siempre.

Y lo hizo con humor. Como a mí se me ocurre decir que se debe hacer siempre. Como he visto que lo han hecho y hacen tantos humoristas que sufrieron uno y más y más *bajones*. Y como a mí se me ocurre también que hacen siempre los humoristas cuando la depresión *no está*, un poco o un bastante como en la canción aquella de mi infancia, que recuerdo así: *Juguemos a la ronda / mientras el lobo está*. Pero que, o tiene ahí metida una enorme elipsis, o qué sé yo, pero que, a mí me parece, debería decir, más bien: *Juguemos a la ronda / mientras el lobo no está...*

Y en aquella novela el desesposado analizó, uno por uno, los efectos secundarios de un antidepresivo llamado Anafranil. Los analizó, por supuesto, literariamente. Y tanto, que el director en Suiza de los laboratorios que fabrican o fabricaban este producto, le replicó varias veces, epistolarmente, desde sus puntos de vista científicos y con toda la información médica sobre aquel producto. Pero resulta que ese científico suizo era tam-

bién un hombre de humor y lo que realmente salió de todo aquel intercambio de cartas fue una verdadera amistad epistolar, en la que ambos reconocían haber aprendido algo y mucho del otro, pero en la que, también, terca pero amistosamente, cada uno se mantuvo en sus trece. El médico suizo en sus *trece* científicas y el desesposado en sus *trece* literarias y en sus *trece* de vida y de auto-observación, que diría el gran Michel de Montaigne.

Y en una sola cosa transaron, de aquélla y de esta crónica de una depresión. De aquellas cartas y de este texto. Los efectos secundarios son dignos de la mejor y la peor literatura, de la misma forma en que varían según quién sea el paciente que se traga el bendito y maldito Anafranil. En fin, como la vida misma. O como diría Oscar Wilde: “El mundo se ha reído siempre de sus propias tragedias, como único medio de soportarlas”. Y muy modestamente, o sólo para encontrarle un punto final a lo que no lo tiene, yo agregaría: “Sí, don Oscar. Tiene usted toda la razón. El mundo se ha reído siempre de sus propias tragedias. Pero lo ha hecho, o lo ha hecho más, o lo ha hecho mejor, a través de sus humoristas, y, en especial, a través de aquellos que conocieron la muerte en vida.

Antonio Fraguas, el gran humorista gráfico y lo que se quiera del madrileño diario *El País*, que de su propia depresión dijo, en una entrevista concedida a la citada revista *Jano*: “Una de las formas de curarse una depresión es verse a uno deprimido y reírse de uno mismo. Recuerdo que tuve una depresión galopante (...) ; “envuelto en mi manta, con barba de 10 días y echado en el sofá, intentaba proyectarme desde arriba, y la verdad es que me daba risa verme tan estúpidamente chafado, y aquello se me fue pasando. Ojo, la depresión no excluye el sentido del humor. Recuerdo que durante ese tiempo jamás dejé de dibujar y de hacer chistes; incluso mi mujer recuerda aquella época y me dice que alguno de los mejores dibujos los hice en aquella época”.

O sea que voy a concluir sin concluir, o mejor dicho repitiendo una palabras de este gran Forges, que dejan esta charla abierta, al hacerla salir de lo íntimo a lo público, o en todo caso arte del buen gobierno, que no la inmundita política. No confundamos, por favor. Y la frase es, and the winner is: Mister Antonio Fraguas, alias Forges, por sus siguientes y ya citadas palabras: “Si todos los estados reconocieran la sonrisa como uno de los derechos humanos, se vería tal vez que al existir este derecho a la sonrisa todos los demás derechos están incluidos”.

No olviden, señoras y señores, que estas son las palabras de un hombre que dijo también que la depresión no lo suelta a uno nunca. De un gran depresivo. Un millón de gracias. □

Margaret Power: *Right-Wing Women in Chile: Feminine Power and the Struggle Against Allende, 1964-1973* (Pennsylvania State University Press, 2002).

EL FEMINISMO EN RETIRADA

A PROPÓSITO DEL LIBRO *RIGHT-WING WOMEN IN CHILE: FEMININE POWER AND THE STRUGGLE AGAINST ALLENDE, 1964-1973*, DE MARGARET POWER

Ana María Stuvén

Escribir sobre feminismo, aunque sea en perspectiva histórica, es materia arriesgada, especialmente en la pluma de una mujer. Invita a formularse preguntas sobre segundas intenciones, reivindicaciones, posiciones antimasculinas, e incluso, libertinajes. Remite también a una oposición con la natural feminidad y con el abandono de la estética y la poética que acompaña al eterno femenino.

La licencia de historiadora y la necesidad de asumir la relevancia de la reflexión sobre los procesos de incorporación de la mujer a la sociedad civil y a la sociedad política animan hacia la incursión dentro de un tema que, talvez por estigmatizado, no ha tenido la merecida atención historiográfica. La reciente publicación de *Right-Wing Women in Chile: Feminine power and the struggle against Allende, 1964-1973* por Margaret Power es un incentivo adicional, ya que al discutir la participación femenina en esos

ANA MARÍA STUVÉN. Doctorada en Historia por la Universidad de Stanford. Profesora/investigadora de los Institutos de Ciencia Política e Historia de la P. Universidad Católica de Chile. Se ha dedicado especialmente a investigar temas de pensamiento y cultura política latinoamericana y chilena. Autora de *La Seducción de Un Orden: Las Elites y la Construcción de Chile en las Polémicas Culturales y Políticas del Siglo XIX* (Santiago: Ediciones Universidad Católica, 2000).

años cruciales de la historia de Chile queda de manifiesto la necesidad de situarla en un contexto de mayor perspectiva histórica, que aclare esa irrupción de las mujeres en la vida pública, así como sus conclusiones sobre el feminismo en Chile.

Indudablemente, parte del problema con la discusión sobre feminismo radica en que el término remite a procesos de distinta índole. Por un lado, se le ha identificado con la lucha militante de mujeres por avanzar su posición en la obtención de derechos en lo político, lo legal y lo económico. En esta acepción, opera como sinónimo de emancipación de la mujer en contra del mundo masculino. En su versión más académica, el feminismo ha sido definido a partir de dos enfoques. El primero es relacional, basado en el género, el cual enfatiza los derechos femeninos con respecto a los de los hombres, por la valorización de la contribución de la mujer desde los roles sociales que ocupa. Esta concepción se apoya en la noción de complementariedad de los sexos, y en una visión que propone una organización social igualitaria. Es la visión que predominó principalmente durante el siglo XIX en la Europa continental. El segundo enfoque proviene de una corriente más liberal, individualista, que enfatiza los conceptos abstractos de derechos humanos individuales, extensivos por su naturaleza, y no por el género, a la mujer. Sus orígenes se remontan a la obra de John Stuart Mill, *La Esclavitud de la Mujer*, publicada en 1869, y a la corriente anglonorteamericana. Aunque el pensamiento feminista francés fue el que más influyó en Chile, Martina Barros de Orrego scandalizó a la sociedad de su época traduciendo la obra de Stuart Mill y publicándola en la *Revista de Santiago*, en 1872¹. El autor inglés sostenía en su obra que el mejor indicador de avance social es la superación de la desigualdad entre hombres y mujeres.

En este ensayo, me propongo contextualizar y discutir algunas de las afirmaciones de Power, desde una reflexión en torno a aspectos de la historia del feminismo y la participación político-social de la mujer en Chile. Para ello, entenderé como feminista a cualquier persona que reconozca la capacidad femenina de interpretar su propia experiencia vital y sus necesidades, así como de reconocer y definir autónomamente los valores propiamente femeninos. Las personas feministas se oponen, por lo tanto, a cualquier forma de injusticia social que afecte a la mujer, y proponen fórmulas para su eliminación, sea en el campo ideológico como programático. Esta definición toma en cuenta tanto la dimensión relacional como la individualista; pretende destruir la jerarquía masculina pero no el dualismo sexual.

¹ Martina Barros de Orrego, *Recuerdos de mi Vida* (1942). La autora relata que las mujeres de su medio la condenaron con la “excomunión” social por su trabajo.

De allí que sea siempre pro-mujer pero no anti-hombre². El feminismo propone un replanteamiento de las formas sociales de ejercicio y reconocimiento del poder en la sociedad como consecuencia de la común humanidad diferenciada. La perspectiva de género hace necesario que sus exponentes reconozcan que las subordinaciones femeninas no pertenecen a su naturaleza; que son construcciones sociales que asignan a la mujer un conjunto de características definidas culturalmente. Tanto la perspectiva de género como el feminismo postulan que sean las propias mujeres quienes definan autónomamente sus propósitos y estrategias, proponiendo así una visión alternativa de futuro³.

Margaret Power realizó una interesante investigación sobre el rol de las mujeres de derecha en la política chilena durante las décadas de 1960 y 1970, desafiando con prudencia a lo políticamente correcto de su país, apelando incluso a fuentes orales con unas 100 entrevistas, especialmente a mujeres actoras o testigos de los fenómenos políticos chilenos de esos años. Digo que desafié con prudencia, pues su texto contiene todas las justificaciones que ella considera moralmente necesarias a fin de no ensuciarse las manos con personas que apoyaron el régimen militar, “que mató, torturó, exilió y empobreció a personas que conozco y quiero así como a cientos de miles de otros chilenos”. Esa precaución, así como su escepticismo para creerles a sus fuentes, convierten el libro en un producto que merece atención por la información que contiene pero que exige una reflexión inserta en un marco de mayor comprensión hacia los desarrollos político-ideológicos del país, así como del proceso de inserción de la mujer en la vida pública en un período que históricamente trasciende tanto el tiempo como los temas que ocupan a la autora⁴. Ella demuestra a lo largo de sus ocho capítulos, primero, que las mujeres fueron actores relevantes para el derrocamiento del gobierno de Salvador Allende; segundo, que, aunque organizados por mujeres de la clase alta chilena, los movimientos femeninos, especialmente Acción Mujeres de Chile y Poder Femenino, incluían a mu-

² Entendemos género como la construcción social diferencial del comportamiento de los sexos, basada en sus diferencias psicológicas o de naturaleza. Hay una extensa bibliografía: Cfr. Gisela Bock, “Women’s History and Gender History. Aspects of an International Debate” (1989); Joan Wallach Scott, *Gender and the Politics of History* (1999); Asunción Lavrín (ed.), *Latin American Women, Historical Perspectives* (1978); Carmen Ramos Escandón (comp.), *Género e Historia* (1992).

³ Gerda Lerner, *The Creation of Feminist Consciousness from the Middle Ages to Eighteenth Century* (1993), p. 274.

⁴ Margaret Power reconoce que le costó creer a Alejandrina Cox, una de sus entrevistadas, que su “sacada de lengua” al general Prats en 1973 fue espontánea, y que ella no favorecía un golpe militar en esos años.

jeros de clase media y baja. El hecho que todos los sectores de mujeres compartieran una visión tradicional de género, en el sentido que los roles femeninos eran principalmente de esposas y madres y que desde esa función o misión ellas defendían la patria, fue un elemento fundamental para la envergadura nacional que asumieron los movimientos políticos femeninos. En tercer lugar, sostiene Margaret Power que a pesar del apoyo de los partidos Nacional, Demócrata Cristiano y del gobierno de los Estados Unidos (al cual denuncia por negarse a facilitarle documentación sobre sus campañas anticomunistas dirigidas a mujeres), los movimientos femeninos actuaron autónomamente. Finalmente, el libro constata la ausencia de una agenda femenina por parte de la izquierda, incluso dirigida a las trabajadoras, la cual se considera uno de los elementos que afectaron el fracaso del experimento socialista. A propósito de esta carencia en el programa electoral de Allende en 1970, se cita a Carmen Gloria Aguayo justificando su posición: “Creíamos que la lucha por la mujer era parte de la lucha por una sociedad mejor”.

Feminismo y socialismo

Respecto de la izquierda chilena, Margaret Power confirma la posición asumida históricamente por los partidos políticos desde el surgimiento del socialismo a comienzos del siglo XX, de cooptar a las mujeres dentro de la lucha de clases, en posiciones de apoyo al mundo masculino, subordinadas a él, y sin reconocimiento de reivindicaciones de género. Sin embargo, a pesar de la falta de apoyo masculino, mujeres de sectores trabajadores hicieron sentir sus propias reivindicaciones, fuera de las estructuras partidarias. En el caso de las obreras, fueron las mujeres por sí mismas quienes desde temprano se organizaron para su mejor sobrevivencia y condiciones de trabajo, alzando sus voces desde fines del siglo XIX. Hasta la década de 1880, ellas permanecían fuera de las sociedades de socorros mutuos que habían surgido desde mediados de siglo, a pesar de no haber una discriminación explícita contra las mujeres en sus estatutos, y de que ellas constituían una importante fuerza de trabajo, especialmente en la industria textil y de la confección. En 1887 se fundó la Sociedad de Obreras No. 1 de Valparaíso, presidida por Micaela Cáceres de Gamboa. En sus estatutos definió su composición obrera, excluyendo de la categoría a las lavanderas y empleadas domésticas. Se proponían “fomentar la instrucción, la moralidad y el bienestar a fin de que (las socias) puedan cooperar eficazmente al bien público”. Estaba explícitamente prohibido tratar cuestiones religiosas

en el seno de la institución⁵. En 1889 se repitió el modelo en Concepción, con la fundación de la Sociedad Ilustración de la Mujer, apoyada por algunos mutualistas que ayudaron a sus esposas y amigas en los trabajos preparativos. Al año siguiente se fundó en Santiago la Sociedad de Socorros Mutuos “Emancipación de la Mujer”, creada por Juana Roldán de Alarcón, con el objetivo de “trabajar por el bienestar, progreso y cultura del estado de la mujer en Chile”⁶. En 1890, en parte por el escándalo que producía el nombre, modificó su denominación a “Protección de la Mujer”. Roldán fue una educadora, cuya misión era fundar escuelas de niñas, aunque también asumía posturas de lucha por la inclusión femenina en las organizaciones obreras, expresadas por ejemplo en su participación en la sociedad “La Fraternidad de Ambos Sexos”, fundada en 1890.

En 1903, Clotilde Ibaceta, presidenta de la Federación Cosmopolita de Obreras en Resistencia de Valparaíso, propició la “emancipación de las obreras para hacer desaparecer el foco de miseria por que atraviesa nuestro sexo... Los fines que perseguimos no sólo son la unión, el ahorro, el mejor y justo salario, sino también la emancipación y engrandecimiento de nuestro sexo”⁷. En el norte fue donde más proliferaron las organizaciones femeninas de lucha⁸. Destacan allí las organizaciones de Antofagasta: Sociedad de Obreras, Instrucción y Socorros Mutuos No. 1, fundada en 1894 por Eloísa Zurita Arriagada, de origen anarquista. Eloísa asumió el liderazgo en la participación chilena en el Congreso Mundial de Librepensadoras de 1906 en Buenos Aires, el cual proponía, entre otros temas, el divorcio absoluto, la supresión de los ejércitos y la igualdad legal de todos los hijos. De todos estos centros, los que alcanzaron mayor relevancia fueron los Centros Belén de Sárraga, fundados por las mujeres librepensadoras de Antofagasta en 1913 y luego en Iquique, Lagunas y Negreiros. Se trataba de instituciones

⁵ Citado en Sergio Grez Toso, *De la ‘Regeneración del Pueblo’ a la Huelga General, Génesis y Evolución Histórica del Movimiento Popular en Chile (1810-1890)* (1997), p. 600.

⁶ *Ibidem*, p. 602.

⁷ Cecilia Salinas, *La Mujer Proletaria: Una Historia por Contar* (1987), p. 34.

⁸ Algunas autoras lo atribuyen al desarrollo económico de la zona, a la concentración poblacional y a la debilidad de lazos tradicionales, religiosos, etc. propios de una sociedad en gestación. Cfr. Edda Gaviola, Ximena Giles, Lorella Lopresti, Claudia Rojas, *Queremos Votar en las Próximas Elecciones. Historia del Movimiento Femenino en Chile, 1913-1952* (1986). También hay que considerar que la participación femenina en la fuerza de trabajo asalariado aumentó significativamente desde fines del siglo XIX, llegando a constituir un tercio de la población económicamente activa en 1920. Cfr. Elizabeth Hutchinson, “El Feminismo en el Movimiento Obrero Chileno: La Emancipación de la Mujer en la Prensa Obrera Feminista, 1905-1908” (1992).

nombradas en honor a la médico andaluza anarquista invitada por Luis Emilio Recabarren a Iquique a dar una serie de conferencias. Belén se paseó por varias otras ciudades, y en 1915 dio conferencias en Santiago, donde fue fuertemente criticada por los sectores más conservadores por su crítica hacia la Iglesia y por las formas que asumía su denuncia de la situación de los trabajadores, si bien sus posturas, acompañando a Recabarren, no incluyeron reivindicaciones de género. Los centros fundados con su nombre realizaban conferencias y promovían la discusión en torno a los problemas de trabajo, religiosos, y ellos sí se ocupaban de la situación de la mujer⁹.

Lo paradójico de estos primeros intentos de organización de sectores trabajadores femeninos en torno a materias de género fue que justamente por la estrecha relación que se creó entre ellos y el Partido Obrero Socialista de Recabarren, las luchas feministas fueron subsumidas en el discurso de la lucha de clases en desmedro del de género. Las presiones para obtener, por ejemplo, el descanso dominical y contra el alcoholismo, tuvieron que darse en forma independiente¹⁰. En este mismo espíritu, en 1905 las mujeres comenzaron a publicar el periódico bimensual *La Alborada*, primer órgano de prensa redactado por una mujer, Carmela Jeria, el cual fue distribuido en las principales ciudades y permaneció hasta 1907. Por sus fuertes vínculos iniciales con el Partido Demócrata, su portada lo presentaba como “publicación social obrera”; no obstante, en 1906 ésta cambia hacia “publicación feminista”. En ese momento, las colaboradoras comenzaron a cuestionar crecientemente el supuesto de que la emancipación de la mujer sería un resultado automático de la lucha obrera, como se lo sugerían los miembros de los partidos políticos de orientación socialista. En esta postura destacó Esther Valdés de Díaz, quien prometió “... un futuro doméstico utópico después de la emancipación de clase y sexo”. Su mención de “lo doméstico”, como un ámbito a depurar para la mujer, indica la valoración de los roles tradicionales femeninos, aún en un contexto que considera las desigualdades de género¹¹. El periódico asumió posiciones crecientemente antimasculinas y las expresó de manera violenta: “Marcharemos resueltas hacia el porvenir por la ruta que nos hemos trazado, miran-

⁹ Luis Vitale, Julia Antivilo, *Belén de Sárraga, Precursora del Feminismo Hispanoamericano* (1999); Cecilia Salinas, *La Mujer Proletaria...* (1987).

¹⁰ Julieta Kirkwood, *Ser Política en Chile. Los Nudos de la Sabiduría Feminista* (1990). La autora sostiene que el feminismo fue moralista respecto del obrero. Edda Gaviola, *et al.*, *Queremos Votar...* (1986), discrepa de la tesis, dando mayor credibilidad a las reivindicaciones feministas.

¹¹ Hutchinson, “El Feminismo en el Movimiento...” (1992), p. 61.

do con desprecio las bravatas de algunos varones que, con harto pesar, ven que la mujer obrera quiere, de una vez por todas, arrojar lejos de sí las crueles ligaduras que la retienen al lado de sus más criminales verdugos: la explotación y el engaño”¹². Esta explotación se expresa en la voluntad femenina de agradar al hombre con sumisión, y en la esclavitud para la que se prestan las mujeres de todas las clases. “La mujer es en la sociedad presente ni más ni menos que un juguete de los caprichos del hombre, e inconsciente se prepara y se adorna para este deshonesto y humillante sacrificio”, escribió una mujer bajo el seudónimo Selva¹³. Con esa afirmación, *La Alborada* reivindica el discurso de género y la lucha emancipadora no sólo contra su situación, sino también contra las representaciones que los hombres han construido de ellas. Por ejemplo, refiriéndose a la virtud, Hermancia Lesguillon afirma: “A la par que los hombres han dictado las leyes, han inventado una palabra con la que nos azotan sin cesar... La virtud para nosotras es la gloria de llegar a no sentir nada humano; es el arte de helar nuestro espíritu... es secuestrarnos y encerrarnos...”¹⁴. El periódico *La Palanca*, órgano de la Asociación de Costureras, sucedió a *La Alborada*, proponiendo “... sostener ruda lucha, dentro de nuestro sexo, para convencernos de lo indigno y despreciable de nuestra condición actual..., debemos emplear toda nuestra energía para llegar a conquistar en la Sociedad el puesto que por derecho natural nos corresponde”¹⁵. A pesar de este discurso, la influencia mencionada de los partidos masculinos y su definición ideológico-política en el contexto de la lucha anticlerical y de clases, no permitió que la emancipación de la mujer constituyese un programa de acción. Por lo tanto, las mujeres de sectores trabajadores continuaron a lo largo del siglo debatiéndose más bien dentro de la estrategia de un feminismo materialista encaminado hacia la superación de la explotación económica¹⁶. La tesis de Margaret Power sobre la inexistencia de reivindicaciones de género en el socialismo chileno de los años de 1970, se confirma, por lo tanto, también históricamente en un período largo. Sin embargo, debe ser matizado con que eso no implicó, necesariamente, ausencia de las mujeres de la arena pública, ni que especialmente las obreras y trabajadoras cercanas al socialismo no sostuvieran posiciones de género desde aproximadamente 1880.

¹² *La Alborada*, N° 19, 11 de noviembre de 1906.

¹³ *Ibidem*, año II, N° 34, 3 de marzo de 1907.

¹⁴ *Ibidem*, año II, N° 34, 3 de marzo de 1907.

¹⁵ “En el Palenque”, *La Palanca*, año I, N° 1, 1 de mayo de 1908.

¹⁶ Asunción Lavrín, *Latin American Women...* (1978).

Feminismo católico

Si nos atenemos a la definición de feminismo propuesta en el comienzo de este trabajo, debemos constatar también que ya desde mediados del siglo XIX las mujeres católicas ingresaron en la esfera pública en defensa de aquello que consideraban propio, y que eran los valores sociales amenazados especialmente por las corrientes secularizadoras propias de la modernidad. En ese ámbito, ellas valorizaron su participación y reivindicaron su derecho a hacerlo. Se trata de sectores femeninos vinculados al republicanismo que inspiró a los grupos llamados liberales y conservadores. Estableciendo una continuidad laxa podemos asociar estos sectores con aquellos que para mediados del siglo XX nuestra autora llama “la derecha”. En estos grupos existe plena coincidencia entre las representaciones del rol social femenino hechas por los hombres, la Iglesia, y los programas educacionales, con los registros de voces femeninas que se conservan. La ausencia inicial de temas de género en estos sectores no implica la ausencia de un feminismo que exige su cuota de participación. Ellas comenzaron desde temprano a hacer oír sus voces en defensa de la Iglesia y del ethos católico, por ejemplo, cuando en 1859 amenazaron con lanzarse a las ruedas del carro que llevara a Monseñor Valdivieso al exilio para impedir su paso, si el gobierno así lo disponía. Posteriormente, cuando en 1865 se discutió en el Congreso la posibilidad de derogar el artículo 5° de la Constitución que establecía la prohibición del culto público de cualquier religión que no fuera la católica, las mujeres se pusieron a la altura del siglo fundando *El Eco de las Señoras de Santiago*, que alcanzó a tener 12 números. No todos sus contemporáneos concordaron con que el periódico salía de plumas femeninas¹⁷. *El Ferrocarril*, más bien liberal, publicó un artículo firmado por “muchas jóvenes santiaguinas” donde sostenían que: “Es olvidar todos los respetos que se le deben a la mujer, suponer, ni por un momento, en labios de ninguna de nuestras dignísimas matronas ni de sus puras hijas, tales conceptos. Estos renglones no los habría trazado ni la mano de una ramera. Habría sido más respetuosa por su sexo que el mal caballero... peor escritor que las aventura a la publicidad premunido por el anónimo”¹⁸. De la polémica en torno a la presencia de mujeres tras las páginas de *El Eco de las Señoras de Santiago* se desprende que la sociedad opinante, a través de sus

¹⁷ Ricardo Donoso lo atribuye a mujeres, e incluso sostiene que recibieron más contribuciones femeninas que las que podían publicar, véase *Las Ideas Políticas en Chile* (1946), pp. 159-161. Más recientemente, lo confirma Erika Maza Valenzuela, “Catolicismo, Anticlericalismo y Extensión del Sufragio a la Mujer en Chile” (1995), p. 195.

¹⁸ *El Ferrocarril*, Santiago, 22 de julio de 1865 s/p.

expresiones en la prensa, asume que las mujeres han ingresado a la opinión pública, independiente de la autoría directa de sus páginas, porque paralelamente a la aceptación o negativa de su existencia como autoras, la discusión acoge la conveniencia y las formas de participación femenina en la sociedad. De allí que el conservador *El Independiente* tome el guante preguntándose: "... por qué se les ha de negar que escriban, que piensen en la cosa pública, que se ocupen de algo más de lo que hasta ahora las ha ocupado?"¹⁹ Sean o no mujeres, su figuración como autoras las lanza a la arena de la opinión pública.

El Eco tiene un claro contenido feminista, dado por la capacidad de las mujeres de establecer la legitimidad de su participación. En su primer número, las mujeres identifican su misión y sus valores. Afirman que ellas "... saben trocar las agujas por la pluma, y se solazan más en escribir en defensa de la ventura de la patria, que en arrancar al piano embriagantes melodías". Por cierto, esta mudanza de roles se justifica en la medida en que ellas extienden su función familiar hacia la sociedad civil como prolongación de ésta: "Hemos pasado largos años cuidando de nuestros esposos, y de nuestros hijos: unos y otros saben que nuestro amor les pertenece; y de ese amor les damos la más flagrante prueba en salir hoy a la defensa de la religión y de la patria; especialmente nuestros hijos..."²⁰. No se vislumbra en los textos de *El Eco* ninguna reivindicación femenina respecto de su rol social. Lo que las mujeres hacen es tan sólo traspasar ese rol de lo privado hacia lo público, justificado por una concepción de la sociedad civil como la reunión de familias bajo una cabeza política. La educadora y defensora de la familia expande su labor legítimamente hacia la defensa de la nación.

La mujer no incluyó los derechos políticos entre sus reivindicaciones respecto de su naturaleza y su acceso a lo público, durante el siglo XIX y comienzos del XX, lo cual se relaciona más que con una debilidad, con la fortaleza de sus vínculos con lo que considera acertadamente las fuentes de poder real en la sociedad del siglo XIX²¹. Aisladas y poco representativas de un movimiento organizado, las escaramuzas sufragistas de mujeres de San Felipe son notables tan sólo como expresión de una posibilidad. Después de haberse intentado inscribir en los Registros, en 1876 en La Serena y

¹⁹ *El Ferrocarril*, 30 de julio de 1865, s/p.

²⁰ *El Eco de las Señoras de Santiago*, N° 1, Santiago: 13 de julio de 1865, s/p.

²¹ Las primeras luchas por los derechos de la mujer en Europa tampoco privilegiaron los derechos políticos, sino los civiles. Incluso George Sand, en Francia, renegó del derecho de las mujeres a las responsabilidades cívicas. Véase Cécile Dauphin, Arlette Farge, "Culture et Pouvoir des Femmes" (1986). En el caso de la Inglaterra victoriana, fue más importante, como causa femenina, el acceso a las universidades que al voto. Véase Peter Gay, *The Bourgeois Experience, Victoria to Freud, Education of the Senses* (1984).

San Felipe, algunas mujeres intentaron derechamente votar por el candidato Vicuña Mackenna. Los artículos del periódico *La Mujer*, dirigido por Lucrecia Undurraga en la década de 1870, son representativos de la pluma femenina: "... No reclamamos fueros civiles para la mujer, aceptamos su incapacidad como ciudadano activo, ...no marcharemos en son de combate contra ningún fin contrario, si es que hay alguno. No queremos tampoco chocar con ningún principio establecido por las leyes, creencias o costumbres"²². En el primer número de *El Eco...*, dirigiéndose a los hombres, las mujeres escriben: "Nos habéis declarado inhábiles para elegir a los representantes de la nación; por muy deshonrosa que sea esta declaración, la aceptamos con gusto; aun la justificamos. Nos habéis excluido de los congresos, y aplaudimos vuestra determinación. Pero no habéis sellado nuestros labios, ni podéis sellarlos"²³. Es la alusión más directa posible al poder de la opinión pública, de la palabra, de la autonomía del pensamiento crítico y sus formas de expresión. Se trata también de un reconocimiento del lugar de poder femenino a través de las funciones que consideran propias, especialmente en la educación de los hijos. Las mujeres comienzan a usar públicamente su razón, como grupo privado, ejerciendo desde allí el poder hacia la esfera pública.

La aceptación de los roles sociales tradicionales no significa necesariamente que la mujer se sienta menoscabada en sus capacidades. Hay una clara refutación hacia quienes oponen una naturaleza femenina a una cultura masculina, estableciendo incompatibilidades entre una y otra. Las mujeres asumen asertivamente su posición valórica, cultural, apoyada en lo que ven como decisiones autónomas y no como producto de sus limitaciones en tanto sexo débil. Por eso defienden su capacidad de participación, protestando desde una perspectiva que podemos considerar de género, contra las limitaciones naturales que se le suponen. En una sección, "carta de Rosa a Luisa", Rosa protestó contra quienes "... ven en las mujeres un mero simulacro de seres racionales, un autómatas doméstico que hable y accione a voluntad del que maneje los resortes...", y denunció ese juicio como proveniente de "... muchos hombres cuerdos, y lo que es peor, algunas señoras"²⁴. En estos escritos se percibe una suerte de feminismo católico, del cumplimiento del deber y la defensa de los valores, el cual se ejerce con todos los resortes intelectuales necesarios equivalentes a los masculinos. De allí que *El Eco* no solamente asuma la defensa de la capacidad femenina

²² En Lucía Santa Cruz, Teresa Pereira, Isabel Zegers, Valeria Maino, *Tres Ensayos sobre la Mujer Chilena* (1978), p. 167.

²³ *El Eco de las Señoras de Santiago*, N° 1, 13 de julio de 1865, s/p.

²⁴ *Ibíd.*, N° 5, Santiago: 10 de agosto de 1865, s/p.

para aprender, sino que sostenga también que la educación la hace próxima al hombre: “Estamos plenamente convencidas de que poseemos la ilustración relativa a nuestra sociedad, y que esa ilustración se difunde proporcionalmente en igual número de hombres y mujeres”²⁵. También afirman las mujeres que, en su campo, “... aunque señoras, hemos estudiado muchas... cuestiones, y entre las que escribimos no faltan quienes saben el latín con bastante perfección”²⁶.

La mujer conservadora defiende la religión porque cree en ella, pero también e importantemente por su alianza con la Iglesia Católica para la preservación del orden social. “La religión ha sido hasta aquí el único dique que ha estado conteniendo turbas ignorantes; pero, si los vínculos religiosos se enervan y se disuelven, ¿quién podrá impedir que se desborde ese torrente? No hay cosa más clara: mientras menos religioso es un pueblo más hay que echar mano a esta fuerza material para reprimirlo” escriben con motivo de las Fiestas Patrias²⁷. De “lava incandescente” califican a las ideas que juzgan irreligiosas por intentar permitir el culto libre de disidentes²⁸. Su consecuencia “... producirá muy pronto los frutos amargos de odios encarnizados, disensiones domésticas, desunión de los matrimonios, desobediencia a los padres, fraudes en los tratos, en una palabra, la relajación en los vínculos sociales”²⁹. Las últimas décadas del siglo XIX estuvieron marcadas por los conflictos de secularización social y laicización de las instituciones, lo cual provocó a las mujeres a comprometerse en defensa del orden establecido. Sin embargo, también hubo manifestaciones de presión femenina por participar y obtener concesiones en la sociedad civil, especialmente en la voz de las educadoras. En 1872, Antonia Tarragó, directora del Colegio Santa Teresa, elevó una solicitud al Consejo Universitario y luego al gobierno solicitando se autorizase a la mujer rendir exámenes válidos para ingresar a la Universidad. No tuvo eco de inmediato, pero 4 años después Isabel Le Brun de Pinochet repitió la solicitud iniciando una polémica que llegó a la Cámara y a la prensa, y culminó en 1877 con el llamado Decreto Amunátegui que les permitió acceder a estudios superiores e integrarse de pleno derecho a la república de las letras, dando también origen a la fundación de liceos de niñas³⁰. La polémica sobre la educación de la mujer puso sobre la mesa el rol social femenino y su derecho a la igualdad.

²⁵ *Ibidem*.

²⁶ *Ibidem*, N° 8, 31 de agosto de 1865, s/p.

²⁷ *Ibidem*, N° 10, 14 de septiembre de 1865, s/p.

²⁸ *Ibidem*.

²⁹ *Ibidem*, N° 8, 31 de agosto de 1865, s/p.

³⁰ El primer liceo de niñas en Santiago, el Liceo 1, se fundó en 1895.

Con motivo de la discusión sobre las llamadas “leyes laicas” que establecían el registro civil, los cementerios laicos y el matrimonio civil, las mujeres nuevamente reaccionaron para evitar el desmoronamiento de su mundo. Señoras de Santiago, Copiapó, Taltal, Coquimbo, San Felipe representaron en cartas enviadas a periódicos su repudio contra la iniciativa legal. Asumiendo como un derecho la participación, en carta entregada al Presidente, y firmada por unas 180 mujeres de clase alta, ellas apelaron a una “... republicana franqueza” que las facultaría para recordar a S.E. su compromiso con la religión³¹. Su rechazo al Presidente Santa María, impulsor de estas leyes, llegó a tal extremo que algunas de ellas se negaban a rezar el Rosario para no tener que pronunciar el “Santa María...”, del Ave María³². En diciembre de 1883, la mayor expresión femenina recogida hasta ese momento en la historia de Chile se dio en una carta enviada al Senado de la República, con la firma de 17 mil mujeres casadas y viudas contra la posible ley de matrimonio civil. Así se demuestra que la defensa del catolicismo por parte de las mujeres pertenecientes a la clase dirigente y educadas al alero de la Iglesia se mantuvo incólume y en pie de guerra. No obstante, a medida que la sociedad se hacía ideológicamente compleja, que la cuestión social golpeaba las puertas de las ciudades, y que sectores medios y populares presionaban por integrar la sociedad civil y obtener derechos, las mujeres respondieron a lo que identificaron como un nuevo llamado hacia un mayor compromiso con las tareas de la beneficencia, influidas también por la Encíclica *Rerum Novarum* de 1891, su llamado a la justicia social, y su denuncia de la amenaza socialista.

Del feminismo de la generosidad a los derechos

El asociacionismo católico para la beneficencia, especialmente La Liga de Damas Chilenas fundada en 1912, fue una instancia de participación femenina, de aprendizaje profesionalizante y de crecientes reivindicaciones de género a medida que su labor se hizo más compleja. Su misión era “cooperar a la acción de la Iglesia y trabajar por el triunfo de la moral y la civilización cristiana”. Otra de sus metas fue la protección al trabajo a través de la creación de sindicatos femeninos como Las Sindicadas de la Aguja; una tienda, y una bolsa de trabajo, todo ello bajo la protección de la

³¹ “La Mujer Chilena”, 29 de julio de 1883, en *Las Reformas Teológicas de 1883 ante el País y la Historia* (1884).

³² Carlos Walker Martínez, *La Historia de la Administración Santa María* (1889), pp. 30-31.

Virgen del Carmen y con el Arzobispo de Santiago como uno de sus Directores. La combinación de un discurso religioso tradicional con la incorporación de mecanismos de defensa de las trabajadoras es una demostración de la evolución de la caridad femenina hacia un discurso y una visión moderna tanto de la economía como del trabajo³³. La agrupación resolvió acudir a la opinión pública a través de la creación del periódico, simbólicamente llamado *El Eco de la Liga de Damas Chilenas* (1912-1915), probablemente evocando al combativo *Eco de las Señoras de Santiago*³⁴. El periódico calificó al movimiento que representaba como de “acción femenina católica”, y negó que se tratase de un “feminismo católico”, pues su propuesta era formar la conciencia social de la mujer, definida como “la defensa de la fe, la lucha contra la inmoralidad y el amparo de los humildes”³⁵. De esta manera, religión, moral y beneficencia pasaban a constituir un ámbito de acción femenina, cuyo alcance cruzaba las fronteras del país y se vinculaba permanentemente con las labores desplegadas por instituciones similares en los Estados Unidos y países europeos. El rechazo al concepto de feminista por algunas autoras se originaba en una comprensión de éste como “... que aleja a la mujer del hogar y del deber y que la convierte en un ser ridículo como la sufragista inglesa, deshonor y desprestigio de nuestro sexo”³⁶. Sin embargo, la Liga aceptó atribuirse un feminismo en pro de la defensa de “lo que nos es querido, y para el bien de nuestras hermanas desheredadas de fortuna”; aquel feminismo que “no saca del orden a la esposa ni quita a la madre su ternura, ni da a la mujer la sed de puestos, de honores o legislación; que quiere... no el bien propio sino el bien de los suyos, de la sociedad y de la Patria”³⁷. A las formas feministas previas se agrega, por lo tanto, un feminismo de la generosidad. A poco andar, esta posición evolucionó hacia posturas de género. En 1918 se organizó en Santiago el Congreso Mariano, último encuentro privilegiado del feminismo católico³⁸. Lo reconoce Rosa Rodríguez de la Sotta, al considerar el encuentro como un “... trascendental paso del feminismo chileno”³⁹. Define el feminismo, como “la doctrina que tiende a hacer reconocer en la mujer los mismos derechos civiles y políticos que posee el hombre. Estudia el rol de la mujer en la sociedad contemporánea, sus derechos respecto al trabajo, a la ins-

³³ Amalia Errázuriz, “La Liga de Damas Chilenas”, p. 317.

³⁴ Le siguió *La Cruzada* (1915-1917).

³⁵ *El Eco de la Liga de Damas*, año II, N° 29, 1 de noviembre de 1913.

³⁶ *Ibidem*, año I, N° 23, 1 de agosto de 1912.

³⁷ *Ibidem*, año I, N° 21, 1 de julio de 1912.

³⁸ Cfr. Mónica Cuevas U., “La Inserción de la Mujer en la Vida Política en Chile: Adela Edwards de Salas y el Asociacionismo Católico, 1912-1935” (2002).

³⁹ Rosa Rodríguez de la Sotta, “El Congreso Mariano” (1918), p. 1.

trucción, su acceso a las profesiones liberales, a la política, etc.”⁴⁰. Sostiene Rodríguez que existe un feminismo cristiano, el cual ha recibido un gran impulso de la Iglesia, y que “... sostiene y alienta... la emancipación efectiva de la persona moral femenina; su desarrollo en valer; su aplicación a todas las ocupaciones que están de acuerdo con sus aptitudes y sus deberes, en toda la amplitud que se lo permiten esos mismos deberes”⁴¹.

El Congreso Mariano dividió sus tareas en los temas de religión, hogar, educación y acción social. La sección dedicada a la acción social acumuló la mayor cantidad de trabajos de perspectiva feminista, y constituye una apreciable muestra de los avances hechos por la mujer en la observación de su condición a partir de sus actividades en la beneficencia y la caridad⁴². Las mujeres que participaron del Congreso Mariano presentaron ponencias sobre los derechos que la mujer consideraba de suyo; sobre su capacidad económica, condiciones de trabajo de las obreras, sindicatos, mutualidad, pornografía, vagabundaje, orfandad y criminalidad infantil, matrimonio civil, beneficencia y otras. Apelando a la igualdad republicana, se proponían modificaciones en su estatus, así como políticas públicas sobre las cuales se sentían con pleno derecho a ejercer su libertad de opinión. Isabel Irrázaval de Pereira abogó por modificaciones al Código Civil respecto de los derechos de la mujer casada y la patria potestad⁴³. María Aldunate sostuvo que: “Si la mujer trabaja tanto como el hombre, si por sí sola es capaz de sostener su hogar, justo será darle derechos para hacer el uso que ella quiera del dinero ganado con su trabajo, y que si puede ahorrar algo, tenga la seguridad de que nadie tomará ese dinero sin previa autorización de ella”. Asimismo defendió el derecho de la mujer a administrar su patrimonio y propuso para ello una modificación al Código Civil. El mutualismo, en perspectiva de la autora, no sólo reporta bienes materiales, sino que se convierte en el aliado pedagógico de las autodenominadas “... católicas de las clases acomodadas...”, que van en auxilio “... de nuestras hermanas de la clase media o de la obrera”, porque enseña orden, economía, constancia, privación voluntaria y otros rasgos que pasan de predicarse como misión divina a virtudes republicanas. El discurso se articuló con una reacción estratégica por parte de las mujeres de adaptación a los tiempos: “... el movimiento social obrero se hará, tarde o temprano; si no se hace con nosotras se hará sin nosotras y contra nosotras...” María Aldunate reconoció

⁴⁰ *Ibíd.*, p. 2.

⁴¹ *Ibíd.*, p. 3.

⁴² Cfr. Robert Curley, “Género y Política en la Acción Social Católica, 1900-1914” (1996).

⁴³ Elvira Lyon de Subercaseaux, “Orientaciones de la Acción Social de la Mujer Cristiana en Chile” (1918), p. 278.

que “... el socialismo nos ganará la delantera”, si no se toma conciencia de los problemas sociales⁴⁴. Ema González, por su parte, reivindicó el derecho de las obreras a formar sindicatos, “... para cumplir mejor, y en forma más llevadera, los deberes que impone el trabajo,... para rechazar las injustas exigencias de los patrones...” La visión cristiana que inspira a la autora demuestra la evolución en el feminismo católico, en la medida en que no sólo insta al cumplimiento de los deberes para con Dios, sino también “... para consigo misma y para con sus compañeras...”⁴⁵. Y si en ese camino, debe luchar por sus derechos, el Cristianismo “... no tiene nada que oponer... Si ella desarrolla su preparación y competencia, si se forma un valor personal o profesional, que pronto será un valor de opinión, no será posible ya discutir por más tiempo, sobre la actuación de la mujer en la vida pública de la nación”⁴⁶. En ese espíritu, sostuvo Rodríguez: “Hacemos votos para que prospere el feminismo inspirado en los nobles fines de levantar el nivel de la personalidad femenina, reconociendo ampliamente sus derechos y procurándole cuanto necesite para el desarrollo y ejercicio de sus facultades, según el rol que debe tener en la sociedad”⁴⁷.

La irrupción de un discurso feminista de derechos en las primeras décadas del siglo XX marca un notable cambio con el feminismo de incorporación de la segunda mitad del siglo XIX. No obstante esta evolución en la reivindicación de su naturaleza, su capacidad intelectual, su poder y la denuncia de la injusticia de sus incapacidades legales, la mujer no dio prioridad a la lucha a favor de derechos ciudadanos paritarios con el hombre entre las consecuencias inmediatas de su nueva postura. El contenido inicial de este feminismo de derechos estuvo marcado por la lucha por los derechos civiles; por mejorar su condición de género y por modificaciones jurídicas que protegieran su condición y reconocieran los roles sociales que desempeñaba la mujer. A pesar de su adhesión de casi medio siglo a las esclavitudes denunciadas por Stuart Mill, Martina Barros sostuvo en su ensayo “Historia del Feminismo y su Desarrollo en Chile”, leído en el acto de su incorporación a la Academia de Letras de la Universidad Católica en 1917, que las mujeres aún creían que: “La superioridad del hombre es indiscutible en todo lo que significa esfuerzo, capacidad mental y resistencia física. La mujer en cambio posee fuerzas morales, jamás superadas por el hombre, que constituyen su valer y su poderío”⁴⁸. Ese predicamento

⁴⁴ María Aldunate Calvo, “Mutualidad, Ahorro, Seguros” (1918), pp. 301-302

⁴⁵ Ema González, “Manera Práctica de Organizar un Sindicato” (1918), p. 292.

⁴⁶ Rosa Rodríguez de la Sotta, “El Congreso Mariano” (1918), p. 5.

⁴⁷ *Ibídem*, p. 6.

⁴⁸ Martina Barros, *Recuerdos de mi Vida* (1942), pp. 296-7.

justifica que las mujeres chilenas no hayan pleiteado los derechos políticos, especialmente el sufragio, hasta la segunda década del siglo. Margaret Power tiene razón cuando sostiene que las mujeres ingresaban a la política en sus roles tradicionales, y no como ciudadanas que tenían los mismos derechos de los hombres a determinar la política de la nación. Ellas se satisfacían con su ciudadanía pasiva, en parte resignadamente, pero también en parte por lo que aparenta ser una conciencia de los lugares de su poder, en el hogar, las instituciones civiles y la opinión pública; la arena propiamente política no constituía un campo donde se debatían sus intereses, y cuando así era, sabían presionar efectivamente para que el mundo masculino no pudiera prescindir plenamente de ellas.

Derechos civiles y políticos: Republicanismo y liberalismo

En 1917, Delia Matte publicó un artículo en la revista *Silueta*, en el cual distinguía dos tipos de feminismo. El que “es político y pelea por la libertad de sufragio”, que es “el feminismo de las misses (activistas) que van a salto de mata, peleando la libertad de sufragio y paladeando el whisky...”. Defiende, en cambio, el feminismo que “tiende a despertar a la mujer al derecho de estudiar, a ilustrarse, ...un feminismo que casi equivale al simple y trascendente concepto de ser, porque quien no se perciba individual y distinto entre la comunidad humana, no tiene derecho a ser”⁴⁹. Este artículo apunta a un hecho fundamental en la historia del feminismo chileno y a su relación con la historia de las ideas políticas en Chile. Desde la Independencia y durante todo el siglo XIX, el ideario que inspiró a la clase dirigente chilena en la construcción del estado y en su imaginario de nación fue el republicanismo inaugurado casi por fuerza sobre una sociedad tradicional⁵⁰. Esa superposición originó que los procesos de equivalencia entre los derechos correspondientes a la sociedad civil y sociedad política en una sociedad liberal fueran pospuestos, predominando una interpretación donde la representación y la participación están en función principalmente de diques contra la tiranía, el gobierno absoluto y la irrupción popular por el ejercicio pleno de la soberanía popular. En este sentido, la defensa del individuo y en consecuencia del derecho individual como anterior a la comunidad y base del derecho es un principio ajeno al pensamiento político

⁴⁹ En Bernardo Subercaseaux, *Genealogía de la Vanguardia en Chile*, s/a, p. 89.

⁵⁰ Benedict Anderson, *Imagined Communities: Reflections on the Origin and Spread of Nationalism* (1994).

que elaboran los publicistas latinoamericanos y a su tradición. Los primeros republicanos entienden la libertad como la ausencia de dominación arbitraria, lo cual no implica necesariamente la plena vigencia de los derechos individuales, mientras el ejercicio del poder tenga limitación y la suspensión de derechos se decreta en función de un bien social⁵¹. Esta visión de la libertad, dependiente del bien común, es la que justifica las exclusiones políticas, y que la atención no esté puesta en el sufragio como expresión de la igualdad política que refleja, sino más bien en las divisiones reales en el seno de la sociedad. No es, por lo tanto, una fuente de poder originaria, ya que no existen las instancias de deliberación necesarias ni ideas en pugna. El sufragio tan solo delimita una clase política que ejerce el poder en nombre de la nación. Es la cultura política de las Luces, que ve en el ejercicio del gobierno por los más capaces la condición de progreso y verdadera libertad⁵². Expresado en otros términos, la sociedad política es el resultado de la deliberación de aquellos que pueden ejercer lo que un articulista de *El Mercurio* llama en 1845, la “razón nacional”, encarnada en la sociedad ilustrada, cuyas discrepancias se limitaban a “cuestiones de marcha”. Ésta se diferencia de la “voluntad nacional”, inclusiva de todos los sectores sociales, y que no encontraba expresión en la soberanía popular⁵³. En definitiva, los mecanismos de exclusión, a través de la propiedad y la educación como sinónimo de racionalidad, no sólo quedaban plasmados en una estructura de clases, sino que cruzaban transversalmente a todos los sectores de la nación en el proceso de integración de la sociedad civil. En este contexto no es de extrañar que la mujer quedara al margen de la participación y la representación políticas. La exclusión de género se encuentra inserta, por lo tanto, en un sistema de exclusiones que ni siquiera se piensa específicamente para la mujer, pues en las libertades e igualdades republicanas el sujeto es masculino. Incluso en el liberalismo, siendo el sujeto individual masculino, no se hace necesario explicitar la exclusión femenina⁵⁴. De allí que, en una sociedad tradicional, donde las diferencias no son asimiladas como producto de la modernidad tolerante y diversa, la mujer fue adquiriendo su identidad como ajena, marginal, a través del discurso de la diferencia sexual, anterior, en tanto natural, a los discursos sobre propiedad y racionalidad. Lo anterior queda de manifiesto con la sorpresa del legislador cuando en 1875 un grupo de mujeres se intentó

⁵¹ Philip Pettit, *Republicanism, una Teoría sobre la Libertad y el Gobierno* (1999).

⁵² Ana María Stiven, “Chile y Argentina: Representación y Prácticas Representativas para un Nuevo Mundo” (2001).

⁵³ *El Mercurio de Valparaíso*, “Espíritu de Reformas”, 10 de noviembre de 1845.

⁵⁴ Joan Wallach Scott, *Only Paradoxes to Offer, French Feminists and the Rights of Man* (1996), p. 5.

inscribir en los registros electorales de San Felipe. El ministro Ignacio Zenteno, interpelado en la Cámara, contestó que efectivamente la Constitución no privaba a las mujeres de ese derecho; no se había hecho necesario explicitar la exclusión. En consecuencia, la ley electoral de 1884 tuvo que explicitarlo, disponiendo en su artículo 40 que las mujeres no podían inscribirse en los registros, aunque reunieran los antecedentes necesarios para ello.

Para su maduración, el proceso de incorporación de la mujer requería de su autoconciencia de individuo. Por lo tanto, cuando Delia Matte, del Club de Señoras, asigna a la mujer a esta categoría, abre el camino para que ella justifique y encuentre sentido a la lucha por los derechos individuales modernos, incluido el sufragio. Sin abandonar su agenda de derechos civiles, enfocada hacia la solución de los problemas correspondientes a la diada madre-hijo, el Partido Cívico Femenino, fundado en 1921, se constituyó en portavoz de ambas inquietudes; de los doce puntos de su programa, el sufragio fue el noveno. Su periódico, *Acción Femenina*, recogió los Estatutos y la Declaración de Principios del Partido, donde éste se declaró feminista, entendido como que “... el verdadero feminismo no desnaturaliza a la mujer, por el contrario, la hace mejor doncella, más noble esposa, más experta madre y sobre todo una excelente ciudadana y una poderosa unidad social para el verdadero progreso de la humanidad”⁵⁵. En uno de sus primeros números, el periódico enfatizó: “el feminismo debe cuidar de no hacer mujeres masculinizadas..., hay que luchar por que su labor sea siempre la adecuada a su naturaleza o dispuesta de manera que no vaya en contra de ella”⁵⁶. Parafraseando a De Tocqueville respecto de la superioridad de la mujer, el periódico reafirma un feminismo asumido “impulsadas por un sentimiento altruista, por generosidad, por dictamen de nuestro corazón”, a diferencia de las antifeministas que actuarían impulsadas por el egoísmo. Por lo tanto, “el triunfo del feminismo no significará la destrucción de la familia”. Por el contrario, obliga a un compromiso profundo “... de nuestro alto y delicado papel en la sociedad: como madre, como educadora o como obrera intelectual o manual”⁵⁷.

Lo anterior no impide que el periódico recoja un interesante y complejo debate en torno al divorcio, y que muchas mujeres lo defiendan y exijan a través de sus páginas. *Acción Femenina* reprodujo en su primer número el mensaje leído por el Presidente Arturo Alessandri en la apertura

⁵⁵ *Acción Femenina*, año I, N° 1.

⁵⁶ *Ibidem*, año 1, N° 6, febrero de 1923.

⁵⁷ *Ibidem*, año 1, N° 4.

de las sesiones ordinarias del Congreso Nacional el 1 de junio de 1922, en el cual se refirió a la situación legal de la mujer, a la necesidad de superar los problemas de los hijos ilegítimos y naturales, a su interés por reivindicar para ellas los derechos ciudadanos, y a su planteamiento respecto de la urgencia en establecer “el divorcio con disolución de vínculo en beneficio de la mujer y por causas justificadas”⁵⁸. En su segundo número, retomando el tema, se recogió una opinión menos radical. Refiriéndose a unas supuestas declaraciones de Esmeralda Centeno de León en Nueva York, sobre que en Chile habría “medio de millón de mujeres que piden divorciarse”, la articulista protestó contra la imagen que proyectaba contra la familia chilena, y afirmó que: “... aceptamos en principio cláusulas legales en nuestros códigos que liberten a la mujer de una unión desagradable...”, lo cual no implicaría a su juicio un problema generalizado, aunque no por eso desmiente la preocupación social por el tema. Ello se confirma con la publicación de la novela feminista *Los Fracasados* de Delia Rouge, donde la autora asumía una posición favorable al divorcio para resolver la vida de los “fracasados” en el matrimonio⁵⁹. La posición oficial del periódico fue de cautela. Por ejemplo, respondiendo a una consulta, el Director sostuvo que, previo al divorcio con disolución de vínculo, sería preferible “preparar a la mujer para que en cualquier momento de su vida esté en condiciones de ganar su subsistencia independientemente”⁶⁰. Hacia el año 24, *Revista Femenina*, sucesora de *Acción Femenina*, se hizo eco de que el divorcio era un tema social llamando a un concurso público con el tema “¿Queremos Divorciarnos?”. El periódico reprodujo dos textos enviados al concurso. En el primero, Elisa Rivera de Hederra asume una posición antidivorcio, argumentando que atenta contra los hijos; que una vida rehecha es inmoral, y que la ley de divorcio ampararía a los hombres que quieren dejar a sus mujeres. La autora abogaba por leyes proteccionistas para la mujer y los hijos, y especialmente por dar a la mujer la instrucción necesaria para ganarse la vida. En el segundo, Delia Rouge, reconoció que el tema estaba en la agenda pública, y que la moneda “tiene dos caras”, las cuales ella intentaba mostrar con objetividad y apelando exclusivamente a argumentos prácticos. No cree que el divorcio traería “desmoralización”, sí reconoce que trae infelicidad para los hijos y los cónyuges, y que deja a la mujer que no sabe ganarse la vida en condiciones económicas desamparadas. Por lo tanto, aunque argumenta a favor de la dictación de la ley, insiste en que ésta

⁵⁸ *Acción Femenina*, año I, N° 1.

⁵⁹ El periódico la comenta en año 1, N° 7.

⁶⁰ *Ibíd.*, año I, N° 3.

debe poner condiciones, entre las cuales menciona la falta de hijos, la locura, la desaparición del cónyuge, y otras. No apela nunca a razones de tipo religioso ni moral.

Cuando en 1935 se reinician las publicaciones del Partido Cívico con la segunda época de *Acción Femenina*, se continúan recogiendo artículos, esta vez todos a favor del divorcio, basados en el argumento de que las leyes deben adecuarse a las necesidades humanas, y que el número de matrimonios fracasados lo justifica. Zulema Arenas Lavín sostiene que el divorcio es "... un factor importantísimo, para el progreso de la moral y por tanto su ley, de imperiosa necesidad, en todo país civilizado". Aún más, argumenta que: "Inutilizar la vida moral y física de la mujer, cuando está apta para desempeñar la sagrada misión (de la maternidad) que las leyes naturales le imponen, es una rebelión injustificada contra la omnipotente Naturaleza, es absurdo, es inhumano e inmoral... pues no es moral la (situación) de apariencias, sino la de verdad"⁶¹. En general, todos los argumentos a favor del divorcio colocan a la mujer como la víctima de los hombres en todo sentido, y, por lo tanto, esta ley protegería a la mujer contra, como dice Mercedes Pinto, las "excesivas libertades" que la sociedad ha dado al hombre⁶².

En su edición de septiembre de 1935, *Acción Femenina* informa del rechazo del proyecto de ley de divorcio y culpa a quienes convirtieron el tema en "un banderín político y religioso, en vez de juzgarlo como un problema social y moral". Evidenciando la evolución en la posición del periódico hacia el apoyo incondicional a la ley de divorcio, el rechazo de la iniciativa es visto como un atentado contra los pobres, pues los ricos "... con su dinero sacan su divorcio adelante donde sea y como sea". Desde la perspectiva de género, el problema se visualiza como un menosprecio hacia la mujer y su capacidad de administrar y rehacer su vida. Sin embargo, tanto este órgano como *Revista Femenina* siempre sostuvieron posiciones conciliadoras entre los roles tradicionales de madre y esposa y el ejercicio de la virtud republicana a través del sufragio y los derechos civiles, incluyendo el divorcio. Una vez conseguido el voto municipal en 1934, el Partido Cívico dejó a sus socias en libertad de acción, en parte porque quería diferenciarse de los partidos políticos masculinos, con una postura crítica hacia ellos por la negativa partidaria a reconocer los problemas de género.

⁶¹ *Ibidem*, año I, N° 12.

⁶² *Ibidem*, año VI, N° 4.

Conclusión

Esta revisión histórica nos permite concluir que si bien el sufragio fue una de las banderas del feminismo chileno, ésta estuvo subsumida dentro de otras reivindicaciones, apareciendo los derechos civiles como una prioridad anterior a los derechos políticos. El feminismo chileno evolucionó desde un reconocimiento de su capacidad de presión hacia la sociedad política y sus lugares de poder en la esfera privada extendida hacia lo público, hacia la búsqueda de beneficios materiales por parte de las trabajadoras y de autoridad y capacidad de ejercicio profesional de la beneficencia por parte de las mujeres acomodadas, sensibilizadas por la “cuestión social” y la necesidad de reconocer en la mujer sus derechos civiles. El concepto de feminismo definido en el contexto de este trabajo ilumina la relación de la mujer con el poder desde los lugares donde ésta lo ejerció autónoma y eficientemente aún al margen de la ciudadanía activa.

Margaret Power sostiene en su libro que las mujeres que se manifestaron públicamente contra el gobierno de Salvador Allende rechazaban tener una postura feminista. Si comprobamos que, desde comienzos del siglo XX, las mujeres obreras se organizaron autónomamente en torno a posiciones de género que podríamos definir como feministas, con agendas públicas y posturas independientes de la militancia política del mundo masculino que las rodeaba; si constatamos que las mujeres de sectores medios y altos se organizaron en torno a la beneficencia, a la cultura, e incluso a la religión bajo el signo del feminismo, ¿qué sucedió entonces en la década del 70 que las mujeres abandonaron estas reivindicaciones? ¿Por qué las organizaciones femeninas rechazaron categóricamente el feminismo como alternativa? Una vía para comprender este fenómeno, la cual escapa completamente a Margaret Power, es recorrer las ideas políticas chilenas y analizar los efectos que produjo el temor al comunismo en las mujeres que no estaban vinculadas ideológicamente con la izquierda partidista, en todos los sectores sociales. El comunismo fue percibido como una amenaza contra los valores familiares, valores que ellas proyectaron hacia la nación en el fragor del combate contra el gobierno de Allende. La década del 60, el hippismo, la revolución cubana, el discurso de la Guerra Fría, calaron hondo en sus imaginarios apocalípticos, fomentando la desarticulación tanto del discurso de clase como del discurso de género. Es más, como sostiene Power, todos los sectores políticos, incluso aquellos que las habían apoyado, se cuidaron de no vincular las manifestaciones femeninas con sus partidos: hacerlo habría implicado que las mujeres tenían motivaciones políticas. El apoliti-

cismo femenino era fundamental para mantenerlas a mano pero distantes de las esferas de influencia; su acción fue una irrupción definitiva e inesperada, pero útil. Lo confirma Sergio Onofre Jarpa, entrevistado por la autora: “las mujeres fueron importantes en términos de publicidad; fueron las acciones masculinas las que hicieron posible el cambio” (p. 183).

A juicio de la autora, la generación que vivió la adolescencia durante el gobierno de la Unidad Popular manifiesta los efectos del régimen militar, en su combinación de ideología del orden y la ley con la defensa de la familia. El feminismo fue reinterpretado como un ataque a la civilización cristiana occidental, lo que convenció a las mujeres que Power llama “de derecha” de la conveniencia de la disolución de todos sus movimientos. Sin embargo, que el voto femenino haya favorecido en un 45,5% a Ricardo Lagos en 1999 en la primera vuelta presidencial y en un 48,7 en la segunda indica que, a pesar del “retiro” voluntario o involuntario que vivió la vida política femenina durante el gobierno militar, las ideas de anticomunismo y temor al caos han cedido junto con las polarizaciones ideológicas que asolaron Chile durante las décadas anteriores. Sin embargo, las luchas feministas que hemos esbozado para el siglo XX y el activismo político que ilustra Margaret Power en su libro no parecen encontrar un eco igualmente combatiente en agendas públicas para temas de género. Llama la atención que en 1920 el divorcio fuera parte de la agenda femenina, defendido en publicaciones periódicas combativas, junto con muchas otras reivindicaciones a favor de la independencia de la mujer. En la década del 2000 esa participación adolece de la pasión que acompañó las luchas femeninas desde 1934, a pesar del reconocimiento oficial al rol de la mujer, del aumento en sus oportunidades de trabajo y de la legitimidad otorgada a sus reivindicaciones.

Right-Wing Women in Chile es un libro importante para la historia reciente chilena. Complementa tantos estudios sobre el derrocamiento de Allende y el gobierno militar donde la mención a las mujeres o la ausencia de ellas evoca más bien las afirmaciones de Sergio Onofre Jarpa. Historiográficamente continúa, aunque en Chile de alguna manera inaugura, una visión que busca a la mujer en los lugares que efectivamente ha ocupado en la historia, para encontrar sus motivaciones y su influencia no solo sobre la vida privada y la cotidianeidad, sino también sobre la vida pública y el poder. Es una invitación exitosa a insertar los estudios de la mujer en la dinámica sociopolítica y cultural chilena.

REFERENCIAS BIBLIOGRÁFICAS

- Acción Femenina*, año I, N° 1; año I N° 3; año I, N° 4; año I, N° 6 (febrero de 1923), año I, N° 7; año I, N° 12; año VI, N° 4.
- Aldunate Calvo, María. "Mutualidad, Ahorro, Seguros". En *Relaciones y Documentos del Congreso Mariano Femenino. Celebrado en Santiago de Chile en Julio de 1918*. Santiago: Escuela Tipográfica La Gratitude Nacional, 1918.
- Anderson, Benedict. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. Londres: Verso, 1994.
- Barros de Orrego, Martina. *Recuerdos de mi Vida*. Santiago: Orbe, 1942.
- Bock, Gisela. "Women's History and Gender History. Aspects of an International Debate". *Gender and History*, Vol. 1, N° 1 (primavera 1989).
- Cuevas U., Mónica. "La Inserción de la Mujer en la Vida Política en Chile: Adela Edwards de Salas y el Asociacionismo Católico, 1912-1935". Tesis inédita de Magister en Ciencia Política, P. Universidad Católica de Chile, Santiago, 2002.
- Curley, Robert. "Género y Política en la Acción Social Católica, 1900-1914". *La Ventana*, Revista de Estudios de Género, N° 4, Guadalajara, 1996.
- Dauphin, Cécile, Arlette Farge. "Culture et Pouvoir des Femmes". *Annales*, marzo-abril, 1986.
- Donoso, Ricardo. *Las Ideas Políticas en Chile*. México: FCE, 1946.
- El Eco de la Liga de Damas*, año I, N° 21, 1 de julio de 1912; año I, N° 23, 1 de agosto de 1912; año II, N° 29, 1 de noviembre de 1913.
- El Eco de las Señoras de Santiago*, N° 1, Santiago, 13 de julio de 1865, s/p; N° 5, 10 de agosto de 1865, s/p; N° 8, 31 de agosto de 1865, s/p; N° 10, 14 de septiembre de 1865, s/p.
- El Ferrocarril*, Santiago: 22 de julio de 1865 s/p; 30 de julio de 1865, s/p.
- El Mercurio de Valparaíso*. "Espíritu de Reformas", 10 de noviembre de 1845.
- Errázuriz, Amalia. "La Liga de Damas Chilenas". En *Relaciones y Documentos del Congreso Mariano Femenino. Celebrado en Santiago de Chile en Julio de 1918*. Santiago: Escuela Tipográfica La Gratitude Nacional, 1918.
- Gaviola, Edda, Ximena Giles, Lorella Lopresti, Claudia Rojas. *Queremos Votar en las Próximas Elecciones, Historia del Movimiento Femenino en Chile, 1913-1952*. Santiago: La Morada et al., 1986.
- Gay, Peter. *The Bourgeois Experience, Victoria to Freud, Education of the Senses*. New York: Oxford University Press, 1984.
- González, Ema. "Manera Práctica de Organizar un Sindicato". En *Relaciones y Documentos del Congreso Mariano Femenino. Celebrado en Santiago de Chile en Julio de 1918*. Santiago: Escuela Tipográfica La Gratitude Nacional, 1918.
- Grez Toso, Sergio. *De la 'Regeneración del Pueblo' a la Huelga General, Génesis y Evolución Histórica del Movimiento Popular en Chile (1810-1890)*. Santiago: Centro de Investigaciones Diego Barros Arana, DIBAM, 1997.
- Hutchinson, Elizabeth. "El Feminismo en el Movimiento Obrero Chileno: La Emancipación de la Mujer en la Prensa Obrera Feminista, 1905-1908". *Revista Proposiciones* N° 21 (1992), Santiago.
- Kirkwood, Julieta. *Ser Política en Chile. Los Nudos de la Sabiduría Feminista*. Santiago: Cuarto Propio, 1990.
- La Alborada*, N° 19, 11 de noviembre de 1906; año II, N° 34, 3 de marzo de 1907.

- “La Mujer Chilena”, 29 de julio de 1883. En Comisión de Católicos, *Las Reformas Teológicas de 1883 ante el País y la Historia*. Santiago: Imprenta Victoria, 1884.
- La Palanca*. “En el Palenque”. Año I, N° 1, 1 de mayo de 1908.
- Lavrín, Asunción (ed.). *Latin American Women, Historical Perspectives*. Connecticut: Greenwood Press, 1978.
- Leguillon, Hermancia. Artículo en *La Alborada*, año II, N° 34, 3 de marzo de 1907.
- Lerner, Gerda. *The Creation of Feminist Consciousness from the Middle Ages to Eighteenth Centur*. Nueva York: Oxford University Press, 1993.
- Lyon de Subercaseaux, Elvira. “Orientaciones de la Acción Social de la Mujer Cristiana en Chile”. En *Relaciones y Documentos del Congreso Mariano Femenino. Celebrado en Santiago de Chile en Julio de 1918*. Santiago: Escuela Tipográfica La Gratitude Nacional, 1918.
- Maza Valenzuela, Erika. “Catolicismo, Anticlericalismo y Extensión del Sufragio a la Mujer en Chile”. *Estudios Públicos*, N° 58 (otoño 1995), Santiago.
- Partido Cívico Femenino. “Estatutos y Declaración de Principios”. En *Acción Femenina*, año I, N° 1.
- Pettit, Philip. *Republicanism, Una Teoría sobre la Libertad y el Gobierno*. Barcelona: Paidós, 1999.
- Ramos Escandón, Carmen (comp.) *Género e Historia*. México: Instituto de Investigaciones Dr. José María Luis de Mora, 1992.
- Rodríguez de la Sota, Rosa. “El Congreso Mariano”. En *Relaciones y Documentos del Congreso Mariano Femenino. Celebrado en Santiago de Chile en Julio de 1918*. Santiago: Escuela Tipográfica La Gratitude Nacional, 1918.
- Salinas, Cecilia. *La Mujer Proletaria: Una Historia por Contar*. Concepción: Editorial LAR, 1987.
- Santa Cruz, Lucía, Teresa Pereira, Isabel Zegers, Valeria Maino. *Tres Ensayos sobre la Mujer Chilena*. Santiago: Editorial Universitaria, 1978.
- Scott, Joan Wallach. *Only Paradoxes to Offer, French Feminists and the Rights of Man*. Cambridge: Harvard University Press, 1996.
- Scott, Joan Wallach. *Gender and the Politics of History*. Nueva York: Columbia University Press, 1999.
- Stuven, Ana María. “Chile y Argentina: Representación y Prácticas Representativas para un Nuevo Mundo”. *Estudios Públicos*, N° 81 (verano 2001).
- Subercaseaux, Bernardo. *Genealogía de la Vanguardia en Chile*. Santiago: Ediciones Facultad de Filosofía y Humanidades, Universidad de Chile, s/a, p. 89.
- Vitale, Luis, Julia Antivilo. *Belén de Sárraga, Precursora del Feminismo Hispanoamericano*. Santiago: Cesoc, 1999.
- Walker Martínez, Carlos. *La Historia de la Administración Santa María*. Santiago: Imprenta El Progreso, 1889. 

**ANTOLOGÍA DEL *DEFENSOR DE LA PAZ*,
DE MARSILIO DE PADUA**

Óscar Godoy Arcaya

INTRODUCCIÓN

Nota biográfica e histórica

Marsilio de Padua nació en una fecha incierta, entre los años 1285 y 1290, en Padua, rica y pujante ciudad del norte de Italia. Pertenecía a una familia de abogados, notarios y jueces. Su padre, Bonmatteo de Mainardini, fue notario de la prestigiosa Universidad de Padua. Su familia había ascendido en la estructura social junto con el auge de las ciudades y la vida burguesa, fenómeno característico del siglo XIV. En la época, se denominaba a la nueva clase burguesa, los *popolanus*, aquellos que son o vienen del pueblo¹, significando su origen en la base inferior de la sociedad estamental de la época. Durante este período se consolida el desarrollo de las ciudades, con autonomía y poder de gobernarse a sí mismas. Ello trajo consigo la aparición y florecimiento de la vida política municipal, animada

ÓSCAR GODOY ARCAJA. Doctor en Filosofía, Universidad Complutense de Madrid. Profesor titular de Teoría Política del Instituto de Ciencia Política de la Pontificia Universidad Católica de Chile y profesor del Institut des Études Politiques de París. Miembro de número de la Academia de Ciencias Sociales, Políticas y Morales del Instituto de Chile. Consejero del Centro de Estudios Públicos.

¹ Jeannine Quillet, Introducción de Marsile de Padoue, *Le Défenseur de la Paix* (París: VRIN, 1968), p. 12.

por un cuerpo de ciudadanos, que participaban en las instituciones políticas de la ciudad. Las prácticas ciudadanas y, como trasfondo, la forma política republicana tuvieron un fuerte influjo en la Italia del siglo XIV. En este contexto, la familia de Marsilio de Padua es una excelente expresión de esa clase ilustrada, con un fuerte influjo cultural, político y económico en ascenso, pero con anclajes en el pueblo, que empieza a asumir un importante protagonismo político.

Acerca de la vida de Marsilio durante su niñez y juventud no hay mucha información. Algo se sabe sobre sus estudios de medicina, filosofía y derecho en la universidad de su ciudad. También existen testimonios de que tanto Marsilio como su familia militaron en el partido de los güelfos, partidarios del Papado en su largo conflicto con el Emperador germano. Este capítulo de la vida de Marsilio es importante, pues, posteriormente, él abandonará esa militancia para pasar al bando del Emperador, el partido de los gibelinos. En esta última fase va a influir mucho su relación amistosa con las poderosas familias gibelinas del norte de Italia, los Visconti de Milán y los della Scala de Verona. Pero, como veremos, mucho más determinante será su querella teológica política con el Papa Juan XXII, instalado en Avignon, bajo la protección del rey de Francia, y su estrecha colaboración con Luis de Baviera.

Las noticias biográficas de Marsilio empiezan a ser más abundantes a partir de su radicación en París, en cuya Universidad hizo estudios de filosofía natural, teología y medicina. Se sabe, además, que ejerció las funciones de rector de la Universidad de París, por algunos meses, entre marzo y mayo del año 1313. También hay noticias de que en 1316 fue testigo de una profesión de fe de un compañero suyo en la Universidad de París. Tres años más tarde, en una carta del Papa Juan XXII, aparece ejerciendo funciones diplomáticas en representación de la familia della Scala.

En realidad, la historia de Marsilio de Padua y su obra *El Defensor de la Paz* (*Defensor Pacis*) empezó en París. Mientras residió en esa ciudad se produjo un gran giro en la orientación de su vida. Pero, no existen datos suficientes para recrear e interpretar cabalmente los pasos que antecedieron a ese cambio radical. En todo caso, los hechos conocidos y su cronología sugieren que Marsilio debió experimentar graves conflictos internos, vacilaciones y contradicciones. Por una parte, hay claros indicios de que en algún momento decidió asumir responsabilidades en el seno de la Iglesia, pues, según consta en documentos, en 1313 postuló a una posición como canónigo en Padua². Marsilio habría gestionado esta petición en Avignon, ese

² Ibídem, p. 15.

mismo año. Sin embargo, parece razonable concluir que ya en 1314 había cambiado su postura frente al Papado. En efecto, durante ese año, ya de retorno a Padua, Marsilio inició la composición de su *El Defensor de la Paz*. El uso de la palabra “moderno”, referido a hechos actuales, acaecidos durante ese año de 1314, hacen verosímil esta versión. Dada la escasez de fuentes, estos datos, inferidos de su obra, constituyen el principal testimonio de su distanciamiento y ruptura con el Papado. *El Defensor de la Paz* fue finalizado el 24 de junio de 1324 en París, ciudad a la cual había vuelto en 1320, y empezó a circular en forma anónima, bajo el seudónimo de “Un hijo de Antenor” (legendario fundador de la ciudad de Padua), sin censura eclesiástica. Los contenidos de algunas tesis del libro fueron declarados heréticos por la Iglesia y su autor condenado a la pena de excomunión. El año 1326, una vez que Marsilio tuvo evidencias suficientes de que su autoría había sido descubierta, abandonó rápidamente París, en compañía de su amigo Jean de Jandun, para refugiarse en Nuremberg, que en ese momento era la sede de la corte de Luis de Baviera.

Marsilio de Padua, al asilarse en Baviera, se puso bajo la protección y al servicio de su príncipe, que desde 1313 mantenía un agudo litigio con la Sede Pontificia. Un conflicto que no era sino un capítulo más en la ya larga historia de confrontaciones entre el Papado y los emperadores del Sacro Imperio romano germánico, que se remontaba a fines del siglo XI y comienzos del XII; jalónada de hechos dramáticos, entre los cuales, en tiempos de Marsilio, se conservaba el vivo recuerdo de las pugnas protagonizadas por el emperador Federico II, llamado Barbarroja, y el Papa Bonifacio VIII. El núcleo duro de esta prolongada querrela giraba en torno a la pretensión del Papado de detentar la plenitud de potestad para transmitir y extinguir, dar y quitar, el poder temporal del Emperador.

Cuando Marsilio se asila en Nuremberg, la disputa entre el Papa Juan XXII y Luis de Baviera había entrado en una fase de alta conflictividad. A la muerte del emperador Enrique VII, en 1313, se abrió el proceso de sucesión imperial, que básicamente consistía en la elección del nuevo emperador. A raíz del empate de votos que se produjo en el colegio electoral entre los dos candidatos, Luis de Baviera y Federico de Austria, ambos se enfrentaron en la batalla de Mühldorf, en 1322. Los dos contendientes le dieron a esta batalla el carácter dirimente de “juicio de Dios”. Luis de Baviera triunfó, y en 1323, la Dieta de Nuremberg lo proclamó emperador electo del Sacro Imperio. En consecuencia, como era la costumbre, Luis de Baviera le pidió a Juan XXII que culminara el proceso de su designación, aprobando su coronación como emperador. Pero Juan XXII se negó. Hay que tomar en cuenta que esa negativa reflejaba el veto de Francia a Luis de

Baviera. El Papa no solamente era francés, además estaba expuesto en Avignon a la presión de la pujante monarquía francesa, y, como príncipe temporal de los Estados pontificios, mantenía una alianza política con Roberto de Anjou, que gobernaba el reino de Nápoles. Por lo demás, tanto el Papa como el rey de Francia habían apoyado la candidatura de Federico de Austria al trono imperial. Esta negativa desencadenó un turbulento período de conflictos entre Luis de Baviera y Juan XXII. Esta crisis dividió a la Iglesia y desencadenó una guerra en el norte de Italia.

La respuesta de Luis de Baviera a Juan XXII fue la invasión de los Estados pontificios, su deposición como Sumo Pontífice de la Iglesia, bajo la acusación de herejía, por sostener una doctrina acerca de la pobreza del sacerdocio contradictoria con la predicación de Cristo, y, en fin, la elección de un nuevo Papa, Nicolás V. La réplica papal fue la excomunión, que se extendía a todos los colaboradores del emperador e incluía la dispensa del juramento de fidelidad de sus súbditos a su persona (lo cual permitía y legitimaba la desobediencia civil y la sedición). Antes de estos acontecimientos, que reseñaremos más adelante, hay que destacar que la corte de Luis de Baviera se transformó en un bullente centro de actividad antipapal, que convocó a la disidencia interna de la Iglesia. Así, además de Marsilio de Padua, se asilaron en ella grandes figuras de la teología, la filosofía y las ciencias, tales como Guillermo de Ockham, Pedro de Cesena, Jean de Jandun y otros. Es obvio que Luis de Baviera recibió un apoyo doctrinal y una legitimidad moral que lo estimularon a radicalizar su postura frente al Papado.

En efecto, a comienzos de 1327, apoyado por sus partidarios del norte de Italia, liderados por los Visconti y los della Scala, Luis invadió ese país a la cabeza de su ejército. A fines de marzo del mismo año fue coronado rey de Lombardía en Milán. Y en octubre arribó a Roma, después de batir la resistencia papal. En esta ciudad fue proclamado emperador por un colegio de ciudadanos y recibió la corona imperial el 27 de enero de 1328. Es desde Roma, entonces, que el emperador excomulga y depone a Juan XXII y hace elegir a Nicolás V como cabeza de la Iglesia. Después de algunos meses en Roma, el emperador Luis se retiró a sus dominios en Alemania, acosado por la resistencia del partido papal. El éxito inicial de la operación militar y política de Luis de Baviera se diluyó rápidamente, pues Juan XXII recobró el dominio sobre Roma y los Estados pontificios y Nicolás V renunció e hizo penitencia en Avignon por su acto de rebeldía e infidelidad al Papado.

Marsilio de Padua acompañó a Luis de Baviera durante la invasión a Italia, colaborando estrechamente en su aventura romana. Se supone que

tuvo un protagonismo central en el diseño intelectual y doctrinario de todas las operaciones de Luis de Baviera contra Juan XXII. En todo caso, el discurso teológico y filosófico sobre los cuales se fundan los actos y decisiones formales del emperador son atribuibles al paduense. Pero se puede decir que esta señalada participación constituye el clímax de sus actividades como asesor del emperador.

Posteriormente, en Baviera, la vida de Marsilio adquiere otro rumbo, alejado del poder y del influjo sobre el emperador. En efecto, a la muerte de Juan XXII, en 1334, el emperador Luis inicia negociaciones diplomáticas con el Papado para superar la crisis y conseguir la aceptación de Avignon de su legitimidad imperial. Esas negociaciones estuvieron a punto de tener éxito, pero en definitiva fracasaron. Luis estuvo a un paso de lograr el reconocimiento de su estatuto imperial, pero una vez más la monarquía francesa impuso su veto irresistible y el curso posterior de sus esfuerzos se vio interrumpido por su muerte. Durante esas negociaciones, Marsilio no solamente fue alejado del círculo de consejeros del príncipe, sino que corrió el grave peligro de ser entregado a la autoridad papal.

Entre sus servicios prestados al emperador, con posterioridad a la invasión de Roma, se cuentan dos informes sobre el matrimonio de su hijo y heredero, Ludovico de Brandeburgo, con Margarita de Maultasch, casada a los once años con Felipe de Valois, con el cual no había consumado la unión conyugal. La Sede Apostólica se había negado a declarar la nulidad de ese matrimonio. Los informes de Marsilio, favorables a la nulidad civil, están incluidos en un opúsculo titulado *Defensor Minor*, que resume las grandes líneas de su obra mayor, el *Defensor Pacis*. También, durante su exilio en Alemania escribió una tercera obra, titulada *De Translatione Imperii*, que trata sobre la transferencia del imperio —la potestad plena imperial— desde los romanos a los griegos y desde éstos a los francos y los alemanes. Transferencia que fue objeto de estudio y debates en la época, pues en ella estaba en juego el establecimiento de la verdad histórica acerca de la función política y jurídica que le correspondía al Papado en la designación del emperador.

Se sabe que la principal actividad de Marsilio de Padua durante sus últimos años fue el ejercicio de la medicina. Murió en Nuremberg el año 1342.

El pensamiento político y teológico-político de Marsilio de Padua

El breve trazado histórico expuesto nos permite situar la discusión intelectual, filosófica y teológica que se desarrolla en la época. La vida de

Marsilio de Padua transcurre durante la transición de Europa a los tiempos modernos. Se puede decir que su obra está entre dos mundos, anunciando, por una parte, ideas que se van a desarrollar en la teoría política moderna, y, por otra, inserto por fuertes relaciones de filiación a la cultura y el pensamiento de su época. La teoría política del siglo XIV refleja perfectamente la situación de cambios que experimenta Europa en ese siglo. Las estructuras políticas, sociales y económicas de la Edad Media, fundadas en el feudalismo y el ejercicio de la soberanía política ejercida *de facto* por el Papado, empiezan a hacer crisis. Ya se dibuja con relativa claridad la aparición de una nueva forma política, el Estado nacional. Distintas fórmulas del naciente derecho constitucional de esos Estados nos indican la gradual aparición y consolidación de poderes soberanos sobre poblaciones y territorios segregados de la Cristiandad. Las nacientes monarquías nacionales, en España, Francia e Inglaterra, empiezan a reclamar el ejercicio de la *plenitudo potestatis* en sus reinos, como lo expresa el axioma político que define al rey nuevo: *rex est imperator in regno suo* (el rey es emperador en su reino). La forma elíptica para atribuir al “rey nacional” aquello que de suyo es el atributo esencial del emperador —el *imperium* o *plenitudo potestatis*— revela que la política transcurre en un período de cambio.

Pero si bien el ideal de gobierno monárquico, con jurisdicción sobre parcialidades de las poblaciones y territorios de Europa, empieza a cobrar forma, la cuestión del momento marsiliano es otra, anterior y preliminar. Esa cuestión, “teóricamente” previa, es la dependencia del poder secular a la plenitud de potestad que el Papado alega detentar. Ésta es la cuestión disputada que enmarca la confrontación con el Papado. Paradójicamente, la disputa no va a dirimirse a favor del Imperio, ya que éste está virtualmente muerto, sino en beneficio de la independencia y autonomía de los Estados nacionales emergentes.

Esta gran polémica, que se despliega a nivel político entre Juan XXII y Luis de Baviera, se desarrolla a nivel intelectual entre la curia pontificia y una pléyade de filósofos, teólogos y juristas, entre los que destaca Marsilio de Padua. Su obra *El Defensor de la Paz* (*Defensor Pacis*) es una de las piezas claves de esa polémica. Esta obra desarrolla un vehementemente alegato contra la doctrina papal de la potestad eclesial o sacerdotal sobre el poder civil. Conviene revisar las ideas centrales que articulan ese alegato, siguiendo el hilo argumental de *El Defensor de la Paz*, cuyos argumentos principales recoge la antología que aquí se ofrece a los lectores.

La clave de bóveda del libro, como lo dice su título, es la paz. Su primer capítulo está consagrado a explicar esta clave. Allí su autor nos dice que la condición necesaria para que la ciudad realice su fin supremo natural

es decir, el bien vivir o felicidad de los ciudadanos, es la tranquilidad o la paz. La concepción agustiniana de la paz que sustenta Marsilio, como “tranquilidad en el orden”³, nos remite al consenso o concordia entre las partes que componen a la ciudad. De este modo, el mayor bien público es la concordia o ajuste armónico entre la pluralidad de elementos que componen a la *civitas*, y su mayor mal, la discordia. Marsilio nos anuncia en este texto, sin nombrarla, que la causa de la intranquilidad o discordia de su época es la intervención del Papado en los asuntos relacionados con el gobierno civil o político⁴.

Marsilio de Padua divide a su obra en tres partes. En la primera parte o *dictio* expone el origen, naturaleza y composición del gobierno civil. La segunda *dictio* trata acerca del gobierno eclesiástico, cuyo propósito es definir los poderes que detenta la Iglesia para la salvación de las almas y sus diferencias con el gobierno civil. La tercera parte, muy breve, recapitula las conclusiones que fluyen de las dos partes anteriores.

En la construcción de los argumentos marsilianos se advierte el fuerte influjo de Aristóteles y Cicerón, por una parte, y, como es claro, de las Sagradas Escrituras, la patrística, la doctrina de los canonistas y la nueva teología y filosofía aristotélico-cristiana surgida en el siglo XIII, por otra.

La autonomía del gobierno secular y civil

En la primera parte, Marsilio sigue muy cercanamente las ideas de Aristóteles expuestas en la *Política*. Hay que tener presente que en el siglo XIII se había producido el redescubrimiento de Aristóteles y su difusión, a la cual contribuyó la traducción al latín de sus obras por Guillermo de Moerbeke. La recepción del filósofo griego ya había generado una abundante producción intelectual. Basta recordar los *Comentarios* de Tomás de Aquino a la *Política* y a la *Ética a Nicómaco*, para no mencionar sino aquellas obras relacionadas con nuestra materia. Pero el trabajo de los “comentarios” o glosa a los textos clásicos no es sino la dimensión más externa de un proceso de apropiación del pensamiento del estagirita. En efecto, la metafísica, la ética y la política de Aristóteles son adoptadas como un punto de partida de una filosofía compatible y en armonía con la teología que emana de las Sagradas Escrituras. La síntesis de Tomás de

³ San Agustín, *La Ciudad de Dios*, XIX, 25; *Obras de San Agustín*, volumen XV (Madrid: Biblioteca de Autores Cristianos, 1978), p. 675.

⁴ Marsilio de Padua, *El Defensor de la Paz* (Madrid: Editorial Tecnos S.A., 1989), I, XIX. De aquí en adelante identificaremos esta obra con las letras *D.P.*

Aquino, admirablemente expuesta en la *Suma de Teología* y la *Suma Contra los Gentiles*, reorienta al pensamiento cristiano medieval, salvando el dilema agustiniano que impedía la armonización entre el mundo natural de la humanidad y el mundo sobrenatural de la salvación. El pensamiento aristotélico puso a disposición de la teología cristiana una concepción de la naturaleza compatible con la doctrina de la gracia. La lectura de *El Defensor de la Paz* evidencia que su autor respira un aire intelectual marcado por la nueva concepción de las relaciones entre ambas esferas. Y aun cuando Marsilio no se refiere explícitamente a Aquino, podemos decir que en sus argumentos reconocemos como trasfondo dos ideas capitales, expresadas por aquél, que sintetizan la verdadera revolución producida por el aristotelismo de la época. La primera es que existe un doble orden de la realidad, dos esferas de realidad (*duplex ordo rerum*), natural y sobrenatural. Esta concepción dual de la realidad le confiere a la naturaleza su propia entidad y autonomía, algo que no se sigue de las premisas teológicas de Agustín de Hipona. La segunda idea nos dice que la naturaleza no es suprimida o anulada por la gracia, sino elevada y perfeccionada por ella (*gratia non tollit naturam sed perficit*). La transposición del *duplex ordo* al ámbito de esa particular realidad que es la política, condujo a Tomás de Aquino a sostener la tesis de que la *civitas* es una *societas perfecta*, o sea, una comunidad suficiente para satisfacer todos los fines humanos y hacer viable el bien vivir o felicidad. Y, del mismo modo, que la Iglesia es también una *societas perfecta*, pero en otra esfera, que es aquella de la salvación; pues, en efecto, ella posee la suficiencia necesaria para que los hombres alcancen la vida eterna y la felicidad sobrenatural, que consiste en la contemplación de Dios.

Como ya sugerimos, no hay evidencias para establecer que Marsilio siguiese a Aquino como un discípulo sigue a su maestro. Pero es razonable pensar que Marsilio recibió el influjo de Aquino, que había muerto en 1274, o sea, cuarenta años antes de la llegada de Marsilio a París (aun cuando la obra del aquinatense no había experimentado la fuerte difusión e irradiación que alcanzó posteriormente). En otras palabras, lo que sostengo es que esas ideas fundantes, más allá de su enunciación en la obra de Aquino, estaban vigentes en tiempos de Marsilio y claramente subyacentes en *El Defensor de la Paz*.

Teoría de la pólis aristotélica como reino

Podemos distinguir en la primera parte de *El Defensor de la Paz* tres grandes nudos argumentales. En primer término, Marsilio desarrolla la con-

cepción de Aristóteles de la ciudad como una realidad natural. El texto marsiliano reconstruye la argumentación de los capítulos 1 y 2 del libro I de la *Política*, acondicionándola a la circunstancia medieval que le sirve de contexto. Así, la *pólis*, como la comunidad que culmina las demandas de la naturaleza humana por adquirir suficiencia (*autarkeía*) para realizar sus fines más altos, adquiere ahora la forma de reino (*regno*). Y las comunidades básicas desde las cuales Aristóteles despliega la emergencia de *pólis*, o sea, la familia (casa) y la aldea, ahora se diversifican en las casas, las aldeas, las provincias y las ciudades, para culminar en el reino. Además, Marsilio sitúa la aparición de la politicidad humana en la aldea. Y para sostener esta tesis, explica que en la casa, el gobierno del más anciano es discrecional, pues no se ajusta a ninguna ley o norma, sino al libre arbitrio del *pater familias*. En cambio, la aldea es ya una comunidad que no tolera la sola autoridad discrecional del más anciano: en ella se impone el gobierno de acuerdo a “lo justo y útil con arreglo a algún ordenamiento racional”⁵, pues sus normas interpretan a todos, emergen de una cierta equidad y son el resultado de un “común dictamen racional”⁶.

A partir del nacimiento de la aldea, la ley es el fundamento de la comunidad. Y, siguiendo esta directriz, Marsilio entiende el reino desde tres ángulos. Por una parte, el reino es una pluralidad de provincias y ciudades bajo un régimen. Enseguida, el reino es una monarquía moderada, es decir, un régimen cuyo gobernante tiene poderes limitados por el imperio de la ley⁷. Y, por último, el reino es una comunidad suficiente o perfecta para la felicidad de los ciudadanos: “finalmente lo necesario para vivir y bien vivir, descubierto por la razón y experiencia de los hombres, llegó a su plenitud y constituyó la comunidad perfecta llamada ciudad”⁸. De este modo, para Marsilio el reino solamente difiere en cantidad respecto de la ciudad-Estado que Aristóteles tuvo a la vista. La ciudad, desde el punto de vista de su extensión ya no es una comunidad de casas y aldeas, sino, además, de provincias y ciudades. No obstante, tanto la *pólis* como el reino se originan en la naturaleza social del ser humano y se definen esencialmente por ser comunidades o sociedades perfectas, dotadas de la plena suficiencia para permitir la realización de todos los fines humanos. A mi juicio éste es el eje de la argumentación de Marsilio acerca de la naturaleza de la sociedad políticamente organizada, su autonomía y la amplitud de su jurisdicción sobre los asuntos seculares.

⁵ D.P., I, II, 4.

⁶ D.P., I, II, 4.

⁷ D.P., I, II, 2.

⁸ D.P., I, III, 5.

El énfasis de la función de la ley como una norma racional que ordena y regula la existencia de la ciudad, en la definición aristotélica de la ciudad, nos permite reconocer la presencia de Cicerón y el influjo de su concepción de la *res publica*, ampliamente generalizada en la Edad Media. En este asunto, y en otros, nuestro autor cruza y hace interactuar entre sí ideas aristotélicas y ciceronianas, siguiendo una tendencia intelectual de la época.

Aristóteles, como se sabe, concibe a la ciudad como una realidad natural. Y aplica a ella las categorías básicas de su ontología, y, en primera línea, aquella de la unidad. La ciudad, en efecto, es una realidad y, como tal, su primera propiedad metafísica es su unidad. Pero la unidad de la *pólis*, no es simple, sino compuesta de partes, es un *syntheton* (compuesto). El estudio de las “partes de la ciudad”, en su empiria, pertenece a la ciencia de la política. Tal es la huella que sigue Marsilio. Ahora bien, esas partes pueden ser estudiadas desde distintas perspectivas. En efecto, son partes de la ciudad las comunidades inmersas en ellas, como la casa, la aldea y la provincia. También son partes de la ciudad los poderes que la gobiernan, como son las instituciones que deliberan y definen las leyes, las que ejecutan acciones gubernativas y las que tienen a su cargo la justicia. Pero Marsilio elige como punto de partida de su propia argumentación la división de los individuos que componen a la ciudad, según las funciones que ejercen al interior de la misma. Y en este punto hace intervenir la concepción de Cicerón sobre los “oficios”. Dice Marsilio que son partes de la ciudad aquellas que pueden discernirse a partir de la división del trabajo, tal como aparece en la *Política*: agricultores, artesanos, obreros, militares, sacerdotes y jueces. La explicación de cada estrato nos remite a *De Officiis* de Cicerón. Pero la línea argumental principal se mantiene apegada a Aristóteles y distingue en esa división otra superior: entre la parte productora (agricultores, artesanos y obreros), que realiza funciones necesarias para la satisfacción del “mero vivir, y la parte gubernativa o directiva (militares, sacerdotes y jueces), que es la causa eficiente de la existencia de la ciudad como una esfera del “bien vivir” o felicidad humana.

En la explicación de la parte gubernativa, Marsilio expone la idea de que en ella la potestad gubernativa, en sentido estricto, está radicada en los que juzgan los asuntos públicos, el legislador y, en forma derivada, el gobernante ejecutivo y los jueces. Este punto es capital porque Marsilio sostiene que el poder coercitivo, cuyo instrumento es la milicia armada, está bajo la autoridad del gobierno y éste, en última instancia, del legislador. Y que, en consecuencia, también el sacerdocio está bajo la suprema y plena autoridad del gobierno civil. Para Marsilio no se trata de la descripción de

una situación histórica, sino de una teoría normativa sobre la naturaleza del poder político. Marsilio de Padua propone esta relación entre el poder civil y sacerdotal plenamente consciente de que el advenimiento del cristianismo es posterior a la vida y teorías de los filósofos antiguos a los cuales recurre para establecer su concepción del reino. Más adelante expondremos sus razones.

La constitución o régimen político y el legislador

El estudio de la parte directiva —que ordena y da unidad a la ciudad— conduce a Marsilio al tema de la constitución o el régimen político. En el tratamiento de esta cuestión Marsilio exhibe un pensamiento novedoso, que en cierta medida prefigura a la teoría democrática moderna. Marsilio recrea la tipología aristotélica de los regímenes políticos (monarquía, aristocracia y *politeia* o república, por un lado, y tiranía, oligarquía y democracia, por otro) y las subdivisiones correspondientes a cada uno⁹. En la circunstancia histórica en que transcurre su vida, la monarquía es el régimen emergente con más difusión y legitimidad. Así, tal como le ocurre a Tomás de Aquino y, en general, a los autores de la época, su atención se concentra en ese régimen. Marsilio distingue varios tipos de monarquía y entre ellas sostiene la superioridad y eminencia de la monarquía regia (*monarchia regale*) en su modalidad “electiva”, que es aquel régimen que además de cumplir el requisito básico de ser un gobierno unipersonal que realiza el bien común, está regulado por la ley, cuenta con el consentimiento de los ciudadanos y, en fin, elige al titular del poder soberano¹⁰. Marsilio desarrolla entonces una teoría de la monarquía electiva, cuyo fundamento directo es el consentimiento de los hombres libres del reino, los ciudadanos.

Ahora bien, esa teoría descansa sobre una concepción acerca del origen de la ley y del gobierno. La ley es anterior al gobernante, por esta razón todos los regímenes rectos contemplan su sujeción al imperio de la ley¹¹. El gobernante, en suma, no es *legibus solutus*, no está “sobre” y libre “de” la ley, como sostenía la doctrina imperial, en Roma, y como va a sustentarse en la teoría política y el derecho constitucional de la monarquía absolutista. Pero si esto es así, ¿quién es el supremo legislador del cual emana la ley?

El legislador, según Marsilio, es el pueblo. Esta tesis le atribuye al pueblo, en lenguaje post-medieval, el principio de la soberanía. Esta tesis

⁹ D.P., I, VIII.

¹⁰ D.P., I, IX.

¹¹ D.P., I, XI.

marsiliana podría inducirnos a atribuirle a Marsilio una modernidad que no tiene. En efecto, el pueblo es el legislador, pero el pueblo no es una agregación de individuos, dotados de iguales derechos políticos, sino un cuerpo, una *corporatio*. Así lo declara Marsilio: “mantenemos de acuerdo con la verdad y la opinión de Aristóteles que el legislador o causa eficiente de la ley es el pueblo o corporación de los ciudadanos”¹². Comprender esta clave es esencial para el lector moderno de *El Defensor de la Paz*. Sin ella, es difícil descifrar el intrincado lenguaje que su autor usa para explicarnos la composición del pueblo y su *modus operandi* para establecer la ley, por una parte, y elegir al soberano, por otra.

El concepto romano de *corporatio* se origina en el derecho privado. La *corporatio* es una *fictio iuris*, una creación ficticia de la ley civil para atribuirle derechos y obligaciones legales a un conjunto de personas naturales vinculadas entre sí en torno a algún fin o utilidad común. La unidad conformada por esa pluralidad de individuos es cualitativamente distinta a ellas y constituye un “cuerpo” artificial, que piensa y opera como si fuera una persona natural. Por esta razón, los individuos que lo componen integran una *universitas* con las características de una *persona ficta*. La *corporatio* como *persona* tenía, para los medievales, dos rasgos adicionales especiales. El primero es su perpetuidad (*perpetuitas*), tenía comienzo pero no fin, gozando así de un estado similar al de la eternidad y el *aeuum*¹³. Y, en segundo término, pertenecen a ella sus miembros, presentes, pasados y futuros; o sea, es una *universitas* transtemporal (*universitas non moritur*). Posteriormente esta doctrina de la corporación tuvo importantes aplicaciones políticas, que van en la dirección de la idea marsiliana de autonomía del gobierno civil. Por ejemplo, los constitucionalistas ingleses del siglo XV la aplicaron a la dinastía reinante. Su efecto fue extraordinario, pues permitió darle estabilidad, continuidad e independencia a la monarquía con sucesión hereditaria, pues impedía el vacío de poder producido a la muerte del rey y eludía la transferencia del poder mediado por la Iglesia¹⁴. Gracias a la concepción y la práctica de la casa reinante o corona como una *corporatio*, el rey nunca moría, porque tenía dos cuerpos, uno físico y otro místico, que pervivía como miembro de la *universitas* familiar o dinástica. Así, la corona recaía inmediatamente, sin vacío de poder que pudiese ser ocupado por la Iglesia, en el miembro de la *corporatio* que por sucesión biológica y legal correspondía. Este ejemplo nos ilustra acerca de la importancia de la

¹² *D.P.*, I, XII, 3.

¹³ Ernst Kantorowicz, *Los Dos Cuerpos del Rey* (Madrid: Alianza Editorial, 1985), pp. 206-297.

¹⁴ *Ibidem*, pp. 297-420.

corporatio como dispositivo conceptual e institucional durante todo este período.

Ahora bien, naturalmente las decisiones y operaciones de la persona ficticia las toman y ejecutan sus miembros activos y presentes (en el caso de una dinastía monárquica, es un individuo, el rey gobernante). Pero ellos “representan” a la universalidad o totalidad de sus miembros (pasados y futuros). En el caso que la corporación tenga muchos miembros presentes y en acto, las decisiones se adoptan a través de la deliberación colectiva y la regla mayoritaria. Pero, en la época, también se concebía a las operaciones de la *corporatio* en forma análoga a las de los seres biológicamente vivientes. Así, por ejemplo, el reino era una *corporatio*, pero, a la vez, un ser viviente, cuya cabeza era el rey, y sus partes inferiores los distintos órganos encargados de cumplir funciones vitales para su existencia.

Marsilio de Padua concibe al pueblo como una *corporatio*. O sea, como una persona ficticia, compuesta de partes. La división del pueblo en partes se entiende de dos modos. En general, son partes de la corporación denominada “pueblo” todos los varones libres¹⁵, habilitados como ciudadanos; o sea, dotados de derechos políticos para participar en las decisiones e instituciones públicas¹⁶. Al cuerpo configurado por todas estas partes ciudadanas, Marsilio llama *universitas civium*. Pero, en particular, la corporación “pueblo” está compuesta por dos partes o clases de ciudadanos: la *maior et melior pars* y el resto; la parte calificada y la parte sin calificaciones. La *maior et melior pars* (también conocida como la *sanior pars*), que los traductores al español llaman sintéticamente “parte prevalente”, denominan al conjunto de los más capaces y mejores, aquellos que tienen “más peso”. Esta distinción supone un criterio selectivo basado en la calidad y no en el número¹⁷: los que son libres, ricos, educados y nobles. Se supone que el resto de la universalidad ciudadana —la parte “no prevalente”— está integrada por los menos dotados y más ignorantes o carentes de educación. Sin la unión de estas dos partes, configurando la unidad del pueblo y su mutua interacción, la parte “no prevalente” solamente sería “plebe” o masa inorgánica, *vulgus*.

La *maior et melior pars* también es denominada *valentior pars*, queriendo significar Marsilio a aquella parte esencialmente comprometida con la preservación y desarrollo de la ciudad. O sea, la parte del pueblo

¹⁵ D.P., I, XII, 4.

¹⁶ Reproduzco la definición de ciudadano que cita Jeannine Quillet, Introducción de Marsile de Padoue, *Le Défenseur de la Paix* (Paris: VRIN, 1968), p. 112: *qui participat principatu iudicativo et constitutivo secundum guberino autarkeiam, id est per se sufficiuntiam*.

¹⁷ Aristóteles, *Política*, VI, 1318 a 4.

compuesta por los ciudadanos virtuosos, animados por el amor al servicio público y la grandeza del reino. Esta parte también recibe la apelación de *principans pars*, por ser ella el núcleo esencial desde el cual despliega sus efectos la causalidad eficiente que anima la existencia, unidad y continuidad de la ciudad. Finalmente hay que decir que esta parte —*maior et mellior, valentior et principans*— “representa” al universo de todos los ciudadanos (*universitas civium*). La idea de representación aparece aquí relacionada con la pertenencia a un mismo cuerpo, del cual se es una parte principal que asume o puede asumir el ejercicio de la voluntad del todo. Esta concepción de la representación no es asimilable a la representación democrática moderna, pues supone que una parte de los ciudadanos son desiguales y que eventualmente les corresponde ejercer mayores poderes que al resto. No obstante, la *valentior pars*, según Marsilio, debe fijarse en cada comunidad según la costumbre o la aplicación de un procedimiento consensuado¹⁸.

Ya lo dijimos, para Marsilio de Padua el legislador es el pueblo. Ahora bien, ¿cómo legisla el pueblo en la práctica? Nuevamente es necesario hacer un rodeo. Marsilio le atribuye al pueblo, sin distinción de sus partes, una facultad general: es capaz de juzgar el contenido de las leyes y discernir si son buenas o malas. Pero, en cambio, no es capaz de crear o “descubrir” la ley, en su acepción sustantiva de “medida de los actos humanos civiles”¹⁹. Bajo estos dos supuestos Marsilio sostiene que el pueblo ejerce su función legisladora en la Asamblea, que es la reunión de la totalidad de los ciudadanos. Ahora bien, el procedimiento que propone es el siguiente, que la totalidad de los ciudadanos o la *maior et mellior pars* (parte prevalente) elija a expertos y sabios en cuestiones legislativas para que estos diseñen el o los proyectos de ley. El segundo trámite es que la *maior et mellior pars* evalúe el proyecto de ley y establezca una propuesta final. Esta última debe ser sometida a la Asamblea de todos los ciudadanos, para que cada miembro individual de ella, sea cual fuera su pertenencia a esta o aquella parte, exprese su juicio, pida enmiendas, por agregación o supresión de las disposiciones que ella contenga, y, en definitiva, apruebe o rechace el proyecto de ley²⁰. Este modelo tiene variantes, pues se puede dar que tanto la Asamblea como la parte principal o prevalente asuman más protagonismo que el indicado en las fases anteriores a la aprobación final de la ley. De allí que Marsilio siempre presente la opción para que el pueblo legisle: o lo hace la totalidad de los ciudadanos o la parte prevalen-

¹⁸ *D.P.*, I, XII, 4.

¹⁹ *D.P.*, I, XII, 2.

²⁰ *D.P.*, I, XII, 5.

te. Esta opción es razonada, pues la elección de una u otra vía, según Marsilio, depende de que se verzan o no los obstáculos que la ignorancia o la carencia de talento puedan poner a la “común utilidad” de todos los ciudadanos. Si esos obstáculos son invencibles, opera la parte prevalente: por una “irracional contestación u oposición no puede impedirse u omitirse lo útil a todos”²¹. Pero en cualquier caso, la decisión legislativa última siempre le corresponde a la universalidad de los ciudadanos (*universitas civium*) reunidos en Asamblea. Así, con cierta ambigüedad, Marsilio intenta sostener una doctrina del legislador fundada en la participación de los ciudadanos y el consentimiento.

Marsilio incluye en su teoría de la ley todas las características establecidas por el derecho romano y la doctrina vigente en la época, como es v.gr. su promulgación. En esta tarea no hace sino seguir el canon. Donde hay un destello de originalidad es en la argumentación que expone para fundamentar la mayor legitimidad de la ley que se origina en el consentimiento popular. En efecto, el mayor peso de su argumento radica en que ese consentimiento es la mejor garantía de la sujeción ciudadana a la ley, pues cada cual, al obedecerla, no hace sino obedecerse a sí mismo. Como ha sido destacado por muchos comentaristas del pensamiento del paduense, esta tesis tiene claras resonancias modernas.

Para la elección del gobernante, en el caso de la monarquía electiva, Marsilio no es muy claro. Aparentemente propone procedimientos similares a los que canalizan la aprobación de la ley. La instauración de un príncipe electivo le corresponde a la totalidad de los ciudadanos (*universitas civium*). En la aplicación de este principio se pueden emplear procedimientos diversos que no deben afectar la naturaleza del consentimiento de los ciudadanos²². Su concepción sobre el origen del poder político supone la supremacía del pueblo, la Asamblea, y la ley sobre el monarca. Y funda el régimen en el consentimiento de los ciudadanos. Pero no olvidemos que la ciudadanía marsiliana no es concebida como un estatuto fundado en la igualdad de derechos políticos, sino en la participación en proporción a las diferencias entre los comunes y los miembros de la parte prevalente o *valentior pars*²³, salvo el caso excepcional de que se superen los obstáculos que la *valentior pars*, como ya vimos, normalmente suple.

²¹ D.P., I, XII, 4.

²² D.P., XV, 1.

²³ Jeannine Quillet, Introducción de Marsile de Padoue, *Le Défenseur de la Paix* (París: VRIN, 1968), p. 38.

La Iglesia y sus poderes de salvación

Las grandes líneas de la concepción política de Marsilio se afirman en la plena suficiencia de la ciudad secular para realizar el bien humano natural. Pero, ¿qué ocurre con esa otra *societas perfecta* cuyo fin es la salvación y la felicidad sobrenatural? La *secunda dictio* de *El Defensor de la Paz*, cuya extensión triplica a la primera, está dedicada a este asunto. Marsilio sostiene una posición extremadamente radical y polémica. Dicho brevemente: no solamente le niega al Papado la suprema y plena potestad sobre el poder secular²⁴, como lo afirmaba la doctrina hierocrática durante siglos, sino que proclama la subordinación del sacerdocio al poder secular civil. Doctrina que ciertamente rompe con la concepción de las dos sociedades perfectas, que afirma Tomás de Aquino.

La Iglesia, según Marsilio, carece de la potestad para transmitir y legitimar el poder secular de los gobernantes civiles, que son los únicos que poseen “jurisdicción coactiva sobre los individuos²⁵. Cristo, según Marsilio, “se quiso excluir a sí mismo y a sus apóstoles del oficio de gobernar o de la jurisdicción contenciosa, o régimen o juicio coactivo cualquiera en este mundo”²⁶. Marsilio, además de negar la potestad política de la Iglesia, sostiene que ella debe renunciar a la acumulación de bienes temporales para la realización de su fin salvífico. Cabe destacar que Marsilio suscribe la doctrina de la pobreza de los franciscanos espiritualistas que fue condenada por Juan XXII. Según esa doctrina, el sacerdocio no puede poseer bienes materiales, ni ejercer el derecho de propiedad. En conformidad a la tesis de Marsilio, la Iglesia no posee plena autonomía para gobernarse a sí misma, ni realizar con independencia su fin esencial. La ausencia de autonomía fundamentalmente consiste en su subordinación al poder secular en asuntos que tocan a la elección del Papa y los obispos, a la convocatoria a Concilio general y la difusión y aplicación de su doctrina en el orden temporal.

La Iglesia, según Marsilio, sólo detenta un poder que el mismo denomina “de llaves”, las que permiten el acceso de los hombres a la salvación y la gloria eterna. Básicamente esas “llaves”, que abren las puertas de la salvación, son el poder de absolver los pecados y de administrar el sacramento de la eucaristía. Pero la Iglesia como sociedad organizada, dotada de estructuras internas de autoridad, normas y facultades autónomas para ejercer jurisdicción sobre sus propios fieles, estaría sometida a la autoridad de las leyes y los gobernantes civiles.

²⁴ D.P., II, I, 3.

²⁵ D.P., II, IV.

²⁶ D.P., II, IV, 4.

Marsilio llega a las conclusiones expuestas por distintas vías argumentales. En primer lugar, esas conclusiones son deducidas del principio que establece la autonomía e independencia del poder civil, considerado en sí mismo, como una nota esencial de la naturaleza de la *civitas* temporal. Enseguida, esas conclusiones expresan una defensa de la autonomía soberana contra la pretensión papal de poseer jurisdicción sobre los gobiernos seculares. En consecuencia, niega la doctrina hierocrática del Papado, tal como está fielmente expuesta en la encíclica *Unam Sanctam* de Bonifacio VIII. En ella se afirma que Cristo le entregó a Pedro y a sus sucesores, es decir, al Papado, una potestad absoluta y plenaria sobre “todo” el orden mundano. La piedra angular de esta tesis es el texto del evangelio de San Mateo, XVI, que dice “Tú eres Pedro y sobre esta piedra construiré mi Iglesia, *todo* lo que atares en este mundo, atado quedará en los cielos, etc.” El término *todo*, que destaco, es susceptible de una significación extensa, que incluye todo lo existente, y, por lo mismo, toda transferencia de poder político a los gobernantes políticos. Sobre ese poder, que algunos rotulan bajo el término “principio petrino”, se fundó durante un largo período de la Edad Media el sacramento de la unción real, a través del cual el sacerdocio entronizaba formalmente al rey, que, desde ese momento, como vicario de Cristo, inauguraba el ejercicio del oficio real. La Iglesia, entonces, era la única “mediadora” de un poder que descendía desde Dios, para depositarse vicariamente en el rey. Contra esta doctrina se levanta la doctrina teocrática de la transmisión directa del poder de Dios al rey, sin la intermediación de los sucesores de Pedro, a la que hemos hecho mención más atrás.

Tratándose de una monarquía electiva, como era el caso del emperador del Sacro Imperio Romano Germánico, se planteaba un problema adicional, que conspiraba con la independencia de los mecanismos de atribución del poder a que aspiraban tanto el imperio como los reinos de la época. Y es que, en el caso del imperio, el Papado alegaba que además de las razones doctrinales expuestas existían otras de carácter histórico. El Papado sostenía que estaba en posesión de derechos constitucionales sobre la elección del emperador.

En el texto *De Translatione Imperii*, Marsilio de Padua recrea el proceso sucesorio a través del cual la plenitud de potestad o *imperium*²⁷ habría pasado de los romanos a los griegos (Imperio Bizantino) y de éstos a los francos (Carlomagno) y enseguida a los germanos (Otón I). La Iglesia afirmaba que a comienzos del siglo IV, el emperador Constantino, una vez converso al cristianismo, había donado el dominio sobre todos los territo-

²⁷ Marsilio of Padua, *Defensor Minor and De Translatione Imperii* (Cambridge University Press, 1993).

rios del imperio al Obispo de Roma, y éste, a su vez, habría renunciado a la administración y supervisión de la contingencia gubernativa de los asuntos seculares, devolviéndoselas al emperador. La llamada “Donación de Constantino” sería, entonces, la fuente de los derechos del Papa, como Obispo de Roma, sobre el imperio. Derechos que el Papado ejercía aprobando la designación del emperador y controlando su comportamiento doctrinario. Así, las transferencias expuestas más arriba se habrían hecho siempre bajo la autoridad y la aprobación del Papado, al atribuir derechos imperiales a una determinada dinastía o persona particular. De este modo, el Papado como garante de la Donación, intervenía decisivamente en la designación del emperador.

Marsilio no pone en cuestión la verdad histórica de la “Donación de Constantino”, que de hecho nunca existió, sino que argumenta que el Papa no puede substituir a la fuente natural del poder civil, que es el pueblo. Marsilio se refugia en el principio general que afirma que solamente hay gobierno legítimo allí donde la comunidad lo ha establecido por consenso. En consecuencia, refuta la conclusión principal que los defensores de la Donación sacan de la misma; a saber, que el Papado es un intermediario del poder imperial. Ni el Obispo de Roma ni ningún sacerdote pueden atribuirse la facultad de transferir la autoridad civil de una persona a otra, sin violar con ello el origen legítimo del poder. Y respecto de las transferencias realizadas en el pasado histórico, Marsilio sostiene que ellas fueron ilegítimas o, en el mejor de los casos, no decisivas (simples sanciones de un hecho irresistible)²⁸.

El “principio petrino” es contestado por Marsilio, fundado en la naturaleza de *societas perfecta* de la ciudad, como ya hemos visto. Pero además es contestado a partir de su propia exégesis del Antiguo y Nuevo Testamento. A mi juicio el peso principal de su posición recae en la respuesta de Cristo a Pilatos: “Mi reino no es de este mundo”, que no es sino una extensión de otro dicho del mismo Cristo, atingente al poder secular: “Dad al César lo que es del César y a Dios lo que es de Dios”. Marsilio, refiriéndose a la respuesta de Cristo a Pilatos, comenta, “mi reino no es de este mundo, es decir, no vine a reinar con un régimen o dominio temporal, del modo como reinan los reyes del mundo...”²⁹ A estos efectos, Marsilio recuerda la orden de Cristo a sus apóstoles de “envainar la espada”, cuando estaba siendo detenido en el bosque de Getsemaní, para significar, por una parte, su obediencia y mansedumbre frente al poder temporal, y, por otra, una expresa prohibición de que el sacerdote use la fuerza y la violencia.

²⁸ D.P., II, XII y XIII.

²⁹ D.P., II, IV.4.

Marsilio, además de establecer la separación del gobierno civil y la misión salvífica de la Iglesia sobre bases puramente racionales, apela al Nuevo Testamento para rechazar, como intrínsecamente contradictorio con aquella misión, el uso del poder coercitivo por parte del sacerdocio. En este punto, el paduano sostiene que el “poder de llaves” le confiere al sacerdote un poder similar al del médico para sanar el cuerpo. El sacerdote, nos dice, ejerce su oficio como aquel “que tiene autoridad para enseñar y obrar como el médico, pero no con potestad coactiva sobre ninguno”³⁰. El médico debe y puede enseñar y sanar sin coerción. Y así debe y puede predicar y salvar el sacerdote. Marsilio es particularmente insistente en que el Evangelio transmite una doctrina de la libertad, incompatible con el uso de la coerción física y moral sobre los hombres.

Una de las tesis más polémicas que plantea Marsilio en esta parte del *Defensor Pacis*, se refiere al primado de Pedro y sus sucesores como cabezas de la Iglesia³¹. La exégesis de los textos sagrados lo conducen a afirmar, en primer término, que Cristo les entregó la esencia del sacerdocio “por igual” a todos los apóstoles y, a través de éstos, a todos los sacerdotes. En segundo lugar, que no le transfirió ningún primado a Pedro y sus sucesores sobre el resto del sacerdocio y los fieles. Alude, como prueba, que no hay rastros en la Iglesia primitiva de ese primado. Y que, más bien, se produjo un fenómeno histórico de acumulación de autoridad y prestigio en el Obispo de Roma, que no tiene una justificación doctrinaria o dogmática en las palabras de Cristo, sino exclusivamente en hechos contingentes. La evidencia histórica que expone Marsilio indicaría que originalmente la Iglesia se gobernó en forma descentralizada, pues cada comunidad cristiana elegía a su Obispo y desarrollaba sus actividades en forma autónoma, sin dependencia de Roma, y sujeta a la autoridad civil en todo aquello que no estuviera relacionada con el culto y no afectara al imperio de la ley civil.

El conciliarismo paduano

Otra cuestión principal que Marsilio aborda en el *Defensor Pacis* se refiere a la definición de la autoridad última para fijar e interpretar los contenidos de la verdad revelada por Dios. Su postura es expuesta en forma análoga a su teoría del legislador civil. Así como el legislador civil supremo es la Asamblea de todos los ciudadanos, el legislador eclesial supremo es el Concilio general de los Obispos, sacerdotes y una representación de los

³⁰ D.P., II, VII, 6.

³¹ D.P. II, XV, XVI y XVII.

fieles (equivalente a la *valentior pars* del pueblo político)³². Marsilio, en forma expresa, niega que el Papa detente la autoridad dogmática de la Iglesia. Ella solamente le correspondería al Concilio, por las razones ya expuestas. Y extrema la dependencia de la Iglesia al gobierno civil al afirmar que a este último le corresponde la convocatoria al Concilio. Como antecedente explica que el Concilio de Nicea fue convocado por el emperador Constantino. Esa postura radical de Marsilio descansa sobre la premisa de que el emperador del Sacro Imperio Romano, como príncipe con plenitud de potestad universal —*curator mundi*—, tiene gobierno sobre los actos civiles a través de los cuales se organiza la vida de la Iglesia.

Como es sabido esta doctrina, en una versión más moderada, fue impulsada durante el siglo XIV y parte del XV por los llamados “conciliaristas”, que alcanzaron el clímax de su influjo en el Concilio de Constanza (1414-18). El conciliarismo adoptó la tesis de que el Concilio general de los creyentes es el órgano supremo de la Iglesia. La soberanía, en otros términos, para tomar las decisiones fundamentales relacionadas con el dogma y el magisterio eclesial estaría radicada en ese cuerpo. En el Concilio de Constanza se hizo un intento por llevar a la esfera constitucional de la Iglesia esta concepción de la supremacía del Concilio. Sus defensores —entre los cuales destacaron Jean Gerson y los cardenales de Ailly y Francisco Zabarella— argumentaron que las definiciones dogmáticas y doctrinales de la Iglesia debían ser tomadas por el Concilio y no por el Papa. El Concilio de Constanza puso fin al gran cisma que afectaba a la Iglesia en ese momento, pero no consiguió que sus decisiones sobre la constitución de la Iglesia prevalecieran.

Nota sobre la selección de textos

Los textos del *Defensor Pacis* escogidos para su inclusión en esta antología provienen de la edición integral, traducida al español y anotada por Luis Martínez Gómez, publicada por Tecnos Editorial S.A., 1989. Agradecemos a la Editorial Tecnos su autorización para difundir la obra de Marsilio de Padua.

El Defensor de la Paz está dividido en tres partes, que a su vez se dividen en capítulos (con numerales romanos). Cada capítulo está subdividido en párrafos (con numerales árabes). Así, por ejemplo, I, XV, 2, nos remite a la primera parte, capítulo quince, párrafo dos.

³² D.P., II, XX, 2.

EL DEFENSOR DE LA PAZ*

PARTE PRIMERA

I. DEL PROPÓSITO SUMARIO DE LO QUE SE VA A TRATAR,
RAZÓN DEL INTENTO Y DIVISIÓN DEL LIBRO

Así, a todo reino ha de ser deseable la tranquilidad, en la cual los pueblos progresan y se asegura la utilidad de las naciones. Pues ésta es la madre hermosa de las buenas artes. Esta, multiplicando, con una sucesión renovada, el género de los mortales, ensancha sus dominios y apura sus costumbres. Y se muestra ignorante de tan altas cosas quien no se ocupó nada en buscarlas¹. Casiodoro, en la primera de sus cartas, en el pasaje citado, expuso las ventajas y frutos de la tranquilidad o paz de los regímenes civiles, queriendo, al poner en éstos, como los más preciados bienes, lo mejor para el hombre, a saber, lo necesario para su vida, que sin la paz y tranquilidad nadie puede conseguir, espolear las voluntades de los hombres para tener paz entre sí, y de ahí la tranquilidad. En lo cual se expresó de conformidad con el dicho del bienaventurado Job, cap. 22: *Ten paz y por ella granjearás excelentes frutos*². Por eso Cristo, hijo de Dios, dispuso que fuese ella la señal y anuncio de su natividad, cuando en la misma quiso que por boca de la milicia celeste se cantase: *Gloria a Dios en las alturas y en la tierra paz a los hombres que él ama*³. Por ello también él mismo deseaba a cada paso la paz a sus discípulos. *Vino Jesús y se puso en medio de los discípulos y dijo: La paz a vosotros*⁴. En Marcos les dice, exhortándolos a tener paz entre ellos mismos: *Haya paz entre vosotros*⁵. Ni sólo que la tuvieran entre ellos, sino les enseñaba que la desearan para los otros. De donde en Mateo: *Y entrando en la casa, saludadla diciendo: La paz a esta casa*⁶. Este es el legado que, echándose encima el tiempo de su pasión y muerte, les dejó en testamento a sus discípulos, diciendo por Juan: *La paz os dejo, mi paz os doy*⁷. Ya su ejemplo, como cumplidos herederos e imitadores suyos, los apóstoles la desearon para aquéllos a los que, por medio de sus cartas, dirigían enseñanzas y amonestaciones, sabiendo que

* Marsilio de Padua, *El Defensor de la Paz*, traducción al español de Luis Martínez Gómez (Madrid: Tecnos Editorial S.A., 1989).

¹ *Cassiod., Mon. Germ. Hisat. Variae*, I, 1, Auct. Ant., XII, 10.

² *Job.*, 22, 21.

³ *Lc.*, 2, 14.

⁴ *Juan*, 20, 19.

⁵ *Mc.*, 9, 50.

⁶ *Mt.*, 10, 12.

⁷ *Juan*, 14, 27.

son excelentes frutos los de la paz, tal como, traído de Job, explicó más largamente Casiodoro.

§ 2. Pero como *los contrarios originan los contrarios*⁸, de la discordia contraria a la tranquilidad sobrevendrán al régimen civil o reino cualquiera resultados e inconvenientes pésimos, como se puede bien ver y es a todos patente en el mismo reino itálico. Porque mientras sus habitantes convivieron pacíficamente, gozaron dulcemente de los frutos de la paz antes nombrados, por ellos y con ellos progresando tan adelante, que llegaron a someter a su imperio toda la tierra habitable. Pero surgida entre ellos la discordia y la pelea, con innumerables trabajos y conflictos, fue vejado su reino y sometido al imperio de naciones extrañas y enemigas. Y así además, por la contienda, quedó tan desgarrado por todas partes, como disuelto, que hizo que cualquiera que quisiera y tuviera algo de poder, encontrara fácil la entrada en él y su ocupación. Ninguna admiración en tal acontecimiento, pues como trae Salustio en su *Yugurta: con la concordia las cosas pequeñas crecen, con la discordia las más grandes se desbaratan*⁹. Por ella los autóctonos, llevados al descarrío del error, se privan de la vida suficiente, sufriendo, en vez de descanso, más pesados trabajos, en vez de la libertad, a la continua el yugo duro de las tiranías, y de tal manera se convirtieron en los más infelices entre los demás que viven civilmente, que su nombre patronímico, que solía dar gloria y seguridad a los que lo invocaban, se ha vuelto en baldón de ignominia y es escarnecido por las otras naciones.

§ 3. En estas tinieblas se precipitan los miserables por la discordia y contienda entre ellos, que, como la enfermedad en el viviente, así se manifiesta ser la depravación del régimen civil. De la cual, aunque son muchas las causas originarias y no poco anudadas entre sí las que suelen habitualmente ocurrir de diversos modos, casi todas las describió el eximio entre todos los filósofos en la ciencia civil; pero hay, sin embargo, fuera de ellas, una singular y muy oculta, que la sufrió y la sufre a la continua el imperio romano, en extremo contagiosa y no menos pronta para introducirse subrepticamente en todas las ciudades y reinos, y a muchos ya intentó con su avidez invadir. Porque esta causa ni Aristóteles ni otro de los filósofos de su tiempo ni anterior pudo ver en su comienzo y en su forma. Pues es y fue ésta una cierta opinión perversa que habremos de explicar después, ocasionalmente tomada de un hecho admirable acaecido después de los tiempos de Aristóteles, realizado por la causa suprema más allá de la posibilidad de

⁸ Arist., *Polit.*, 1. 5, c. 8; 1307 b 29. Seguimos la distribución de libros de la ed. I. Bekker y de la trad. de W. Ross.

⁹ Salust., *Yugurta*, c. 10, 6.

la naturaleza inferior y más allá de las acciones que suelen ejercer las causas en las cosas. Esta, pues, con su sofisma y cara de honestidad y de conveniencia, existió para mal del género humano y es apta para traer mal intolerable a toda ciudad y patria si no se le cierra el paso.

§ 4. Son, pues, como dijimos, excelentes los frutos de la paz o la tranquilidad, y de la contraria discordia intolerables los males: por lo cual debemos desear la paz, buscarla si no la tenemos, encontrada guardarla, y con todo el empeño rechazar la contraria discordia. A ella cada uno de los hermanos ha de contribuir, y mucho más las agrupaciones y comunidades entre sí, tanto por el afecto de la caridad como por el vínculo o el derecho de la sociedad humana. A ella también amonesta Platón, según Tulio, en el *De los deberes*, libro primero, cuando dice: *No hemos nacido sólo para nosotros, parte de nuestro existir lo reclama la patria, parte los amigos*¹⁰. A la cual sentencia añade Tulio: *Y, como place a los estoicos, lo que se engendra en la tierra, se crea para la utilidad de los hombres, y los hombres son engendrados para los hombres, y en esto hemos de tomar por guía a la naturaleza, aportar lo nuestro a las necesidades comunes*. Y porque sería de común utilidad, no pequeña, más aún, de necesidad, desennascarar el sofisma de la dicha singular causa de las contiendas, y para los reinos y comunidades amenaza de males no pequeños, cada uno debe poner vigilante y diligente cuidado, queriendo y pudiendo mirar a lo útil. Porque de no saberlo, no se puede evitar esta peste ni cortar del todo su efecto funesto para los reinos y las sociedades civiles.

§ 5. Ni debe nadie ser negligente en este cuidado por temor o por pereza o por cualquier otro motivo inspirado por el espíritu maligno. Pues en la segunda a Timoteo, 1º: *No nos dio Dios espíritu de temor, sino de fortaleza y caridad*; de fortaleza y caridad para propagar la verdad, de donde añade allí el mismo apóstol: *No te avergüences del testimonio del Señor nuestro*¹¹. Y este fue el testimonio de la verdad, para dar el cual Cristo dijo haber venido a este mundo, como en Juan, 18, dice: *Y yo para esto nací y para esto vine al mundo, para dar testimonio a la verdad*¹²; la que lleva al género humano a la eterna salvación. A su ejemplo, pues, a enseñar la verdad, con la cual la sobredicha peste de los regímenes civiles puede caer, sobre todo en el pueblo cristiano; a la verdad, digo, que lleva a la salvación de la vida civil, y es no menos provechosa para la eterna, aquél mayormente está obligado a mirar con empeño, al que el dador de las gracias más inspiró en la inteligencia de estas cosas; y más gravemente

¹⁰ Cic., *De officiis*, I, 7.

¹¹ II Tim., I, 8-9.

¹² Juan, 18, 37.

peca, ingrato, sabiendo y pudiendo esto, si lo omite, según testifica Santiago en su epístola canónica, 4°, cuando dice: *En el que sabe hacer este bien y no lo hace, es pecado*¹³. Pues de otra manera no se quebrantará completamente esta malicia común enemiga del género humano, ni se secarán los frutos perniciosos que hasta ahora ha producido, si antes no se la pone al descubierto y se impugna la maldad de su causa o raíz. Porque por este camino y no por otro puede abrirse paso franco al poder coactivo de los príncipes para derribar finalmente a los perversos fautores y pertinaces defensores de esta maldad.

§ 6. Atendiendo y siguiendo las amonestaciones de Cristo, de los santos y de los filósofos, yo, hombre antenóríde [paduano], si se me ha concedido alguna gracia en la inteligencia de estas cosas, con espíritu de fe venido de arriba, según el testimonio de Santiago en su epístola canónica, cap. 1º: *Toda dádiva óptima y todo don perfecto desciende de arriba del padre de las luces*¹⁴, por reverencia al que lo da, por amor a la verdad y su propagación, por el fervor de la caridad para con la patria y con los hermanos, por compasión y remedio de los oprimidos, para disuadir a los opresores con enmienda de su error, para inquietar a los que lo permiten debiendo ponerle remedio, y mirándote a ti como singular ministro de Dios que ha de llevar esta obra al término que ella espera obtener de fuera, para ti, inclitísimo Luis, emperador de romanos, a quien por la antigüedad de la sangre y herencia y no menos por la singular y heroica condición y preclara virtud se le ha impreso y consolidado un celo de extirpar las herejías, de defender la católica verdad y fomentar y guardar toda otra sana doctrina, cortar los vicios, propagar el cultivo de la virtud, extinguir las contiendas, difundir y promover la paz y la tranquilidad por doquiera, puse por escrito la suma de las reflexiones que siguen, después de un tiempo de diligente y atenta investigación, pensando que de ellas se seguiría alguna ayuda a tu vigilante majestad, ocupado en proveer a los fallos descritos y a los que puedan ocurrir y a las demás utilidades públicas.

§ 7. Es, pues, mi propósito, con la ayuda de Dios, poner en claro esta singular causa de la contienda. Porque tratar del número y naturaleza de aquellas que asignó Aristóteles fuera excesivo. Pero de esta que ni Aristóteles alcanzó a ver ni después de él ninguno, aunque lo pudiera, tomó en consideración, queremos de tal manera quitar el velo que la oculta, que en adelante pueda ser fácilmente eliminada de los reinos y de todas las comunidades civiles, y una vez eliminada, puedan más seguramente los atentos gobernantes y los súbditos vivir con tranquilidad, que era el anhela-

¹³ *Sant.*, 4, 17.

¹⁴ *Sant.*, 1, 17.

do propósito al comienzo de esta obra, necesario a los que han de gozar de la felicidad civil, la que en este mundo aparece como lo mejor y más deseado de todo lo posible al hombre, y el supremo fin de las acciones humanas.

§ 8. Así que dividiré en tres Partes el asunto propuesto por mí. En la primera demostraré el intento por las vías encontradas por el ingenio humano, con proposiciones firmes, de por sí evidentes a toda mente de naturaleza no corrompida por la costumbre o por la pasión descarriada. En la segunda, lo que creí haber demostrado, lo confirmaré con testimonios de la verdad de validez eterna, con las autoridades de los santos intérpretes de ellas, y también de otros probados doctores dentro de la fe cristiana; de modo que el libro este se mantenga firme por sí mismo, sin necesidad de apoyos probatorios foráneos. De aquí también impugnaré las falsedades opuestas a mis asertos y desharé los sofismas que los contradicen con sus lucubraciones. En la tercera inferiré algunas conclusiones o sentencias previstas de certeza evidente, deducidas de lo antes declarado y máximamente útiles para ser observadas, tanto por los gobernantes, como por los súbditos. [...]

II. DE LAS PRIMERAS CUESTIONES DE ESTE LIBRO; DISTINCIÓN Y DEFINICIÓN DE SIGNIFICADOS DE LA PALABRA REINO

Entrando, pues, en nuestro propósito, queremos primero declarar qué sea tranquilidad e intranquilidad del reino o de la ciudad, y de esto, primero la tranquilidad, pues no aclarada ésta, por fuerza se ignorará en qué consiste la intranquilidad, y como ambas parecen ser disposiciones de la ciudad o del reino, lo que se suponía en Casiodoro, primeramente vamos en consecuencia a declarar qué sea y por qué exista el reino o la ciudad, de lo que también aparecerán mejor los rasgos de la tranquilidad y de su opuesto.

§ 2. Queriendo, pues, según el antedicho orden, describir la tranquilidad de la ciudad o del reino, no sea que por la multiplicidad de nombres se origine en este asunto alguna ambigüedad, es conveniente no olvidar que la palabra esta, *reino*, en una de sus significaciones importa pluralidad de ciudades o provincias comprendidas bajo un régimen; según esta acepción no difiere reino de ciudad en la especie de la política, sino más bien en la cantidad. Pero en otra acepción este nombre *reino* significa una especie de política o régimen templado al que llama Aristóteles *monarquía templada*. Y así lo mismo puede existir un reino en una única ciudad que en muchas, al modo como se dio en el comienzo de las comunidades civiles en que, por

la mayor parte, había un rey en una única ciudad. La tercera significación de este nombre y más conocida es la que junta la primera y la segunda. La cuarta es algo común a toda especie de régimen templado, sea en una o en muchas ciudades, y según esta significación la tomó Casiodoro en la proposición que pusimos al comienzo de este libro, y de la cual usaremos también cuando tomemos este nombre en las resoluciones de las cuestiones.

§ 3. Habiendo, pues, de describir la tranquilidad y su opuesto, asentemos con Aristóteles, 1º y 5º de su *Política*, caps. 2º y 3º, que la ciudad es como una naturaleza animada o animal¹. Porque como el animal bien constituido según su naturaleza se compone de ciertas partes ordenadas entre sí con proporción, y con sus funciones combinadas entre sí y en orden al todo, así la ciudad se forma de determinadas partes cuando está bien constituida según razón. Cual es, pues, la relación del animal y sus partes a la salud, tal parece ser la relación del reino o de la ciudad a la tranquilidad. Y el apoyo para esta ilación lo podemos tomar de lo que todos entienden por una y otra. Pues entienden que la salud es la mejor disposición del animal según su naturaleza, y del mismo modo que la tranquilidad es la disposición óptima de la ciudad instituida según razón. Y la salud, como dicen los más peritos entre los médicos al describirla, es la disposición buena del animal, en la cual cada uno de sus miembros puede ejercitar perfectamente las acciones propias de su naturaleza; y según esta analogía la tranquilidad será la buena disposición de la ciudad o del reino, en la cual cada una de sus partes puede realizar perfectamente las operaciones convenientes a su naturaleza según la razón y su constitución, y como el que bien define, significa igualmente su contrario, la intranquilidad será la mala disposición de la ciudad o del reino, como la enfermedad del animal, por la cual están impedidos todos o algunos de sus miembros para hacer sus operaciones propias, o tomados aparte absolutamente, o en su conjunto y funcionamiento total.

De la tranquilidad, pues, y de su opuesto la intranquilidad quede así dicho por modo de comparación.

III. DEL ORIGEN DE LA COMUNIDAD CIVIL

Como ya dejamos dicho que la tranquilidad es la buena disposición de la ciudad referida a la obra de sus partes componentes, convendrá consecuentemente considerar qué sea y por qué exista la ciudad mirada en sí misma, qué y cuántas sean sus partes primordiales; después la acción correspondiente a cada una de ellas y también sus causas y orden entre ellas;

¹ *Polit.*, l. 1, c. 5; 1254 a 31-39; l. 5, c. 3, 1302 b 33; l. 4, c. 4, 1290 b34.

pues éstas son premisas para la determinación perfecta de la tranquilidad y su opuesto.

§ 2. Pero antes de que tratemos de la ciudad, de sus especies o modos que hacen lo que es una comunidad perfecta, debemos primero declarar el origen de las comunidades civiles y de sus regímenes y modos de vivir. Desde los cuales, como desde lo imperfecto, progresaron los hombres a las comunidades perfectas, regímenes y modos de vivir en ellas. Porque el paso y el orden de la naturaleza y del arte imitador de ella es siempre de lo menos perfecto a lo más perfecto. Ni los hombres saben de alguna cosa sino cuando conocen sus *primeras causas y primeros principios hasta llegar a los elementos*¹.

§ 3. Entrando, pues, de este modo en el tema, conviene que no se nos oculte que las comunidades civiles, según las diversas regiones y tiempos, comenzaron de lo pequeño y poco a poco, tomando incremento, finalmente llegaron a la consumación, como dijimos que acaece en toda acción de la naturaleza o del arte. Porque la primera y más reducida de las uniones humanas, de la que las otras a su vez provinieron, fue la del varón y la hembra, como dice el eximio entre los filósofos en el primero de la *Política*, cap. 1^o² y aparece esto más claramente en la *Económica*³ del mismo. Porque de ésta se propagaron los hombres que primero llenaron una casa; a partir de ella se hicieron luego más amplias reuniones, y tan ancha propagación de los hombres, que no les bastase una casa única, sino que hizo falta hacer muchas casas, cuya pluralidad se denominó aldea o poblado; y ésta fue la primera comunidad, como se escribe también allí donde hemos citado antes.

§ 4. Pero mientras los hombres estuvieron en una única casa, los actos todos, que llamaremos después con mayor propiedad civiles, se regulaban por el más anciano de ellos, como el más sensato, sin ley todavía ni costumbre alguna, puesto que aún no se habían podido descubrir. Ni sólo los hombres todos de una única casa se regían de este modo, sino también de modo más o menos parecido la comunidad llamada aldea, aunque con algunas diferencias. Porque si al *pater familias* de la casa única le era dado el perdonar y el castigar los atropellos domésticos según su parecer y omnímoda voluntad, no le hubiera sido esto permitido, sin embargo, presidiendo la primera comunidad llamada aldea. Porque en ésta convino que el anciano dispusiera lo justo y lo útil con arreglo a alguna ordenación racional y ley cuasi natural, por el hecho de que así parecía a todos convenir, con una

¹ *Phys.*, I, 1, c. 1; 184 a 10 ss.

² *Polit.*, I, 1, c. 2; 1252 a 26.

³ Seudo-Arist., *Oeconom*, I. 1, c. 3; 1343 b 8.

cierta equidad, sin aparato aún de gran averiguación, sólo con el común dictamen de la razón y cierta exigencia de la sociedad humana.

La razón de la diferencia de régimen en una casa única y en la aldea es y fue que si en la única y primera organización o familia doméstica el hermano mataba u ofendía al hermano, le era lícito al *pater familias*, sin riesgo, si así lo quería, no castigar al delincuente con el último suplicio; ya porque sólo al padre perdonador parecía habersele hecho la injuria, ya por la penuria de los hombres, ya también porque menor daño y tristeza significaba carecer de un hijo que de los dos; lo cual parece que es lo que hizo el primer padre nuestro Adán cuando su primogénito Caín mató a su hermano Abel. Pues no pertenece propiamente la justicia civil al padre respecto del hijo, como se escribe en el 4º de la *Ética*, en el tratado de la justicia que allí se contiene. Pero en la comunidad primitiva, aldea o vecindario, no fue ni sería lícito obrar así por la disparidad señalada; más aún, si no se hubiera hecho o no se hiciera por el anciano la venganza o reparación equilibradora de los daños inferidos, se habría seguido o se seguiría de allí la lucha o disensión entre los vecinos.

Multiplicándose las aldeas y crecida la comunidad, lo que tuvo que suceder aumentando la propagación humana, todavía se regían por uno, o porque no había muchos hombres prudentes, o por cualquier otra causa, como se escribe en el 3º de la *Política*, cap. 9º⁴, pero siempre por aquel que era tenido, como más anciano, por el mejor, aunque ya con normas menos imperfectas que aquéllas con las que se ordenaban las cosas en la simple aldea o vecindario. Ni tuvieron las comunidades aquellas primitivas tal distinción y organización de las partes componentes, ni el conjunto de las artes y ordenamientos necesarios para vivir, cuantos después se fueron encontrando progresivamente en las comunidades perfectas. Pues alguna vez fue el mismo hombre príncipe y agricultor o pastor de ovejas, como Abraham y otros muchos después de él, lo cual en las comunidades perfectas ni conviene ni sería permitido.

§ 5. Crecidas éstas con el tiempo, creció la experiencia de los hombres, se inventaron las artes y las reglamentaciones y fueron más perfectos los modos de vivir y más distinguidas entre sí las partes componentes de las comunidades. Finalmente, lo necesario para vivir y bien vivir descubierto por la razón y la experiencia de los hombres, llegó a plenitud y se constituyó la comunidad perfecta llamada ciudad, con distinción de sus partes, cuya naturaleza y peculiaridades vamos a declarar.

Baste lo dicho sobre el origen de la comunidad civil.

⁴ *Polit.*, l. 3, c. 14-15; 1285 a 2 - 1286 b.

IV. DE LA CAUSA FINAL DE LA CIUDAD, DISTINCIÓN DE SUS OBJETIVOS
Y DE SUS PARTES EN GENERAL

Es, pues, la *ciudad*, según Aristóteles, 1º de la *Política*, cap. 1º: *una comunidad perfecta, que llena por sí todos los requisitos de suficiencia, como es consecuente decir, creada ciertamente para vivir y persistiendo para bien vivir*¹. Y lo que dice Aristóteles: *creada para vivir y persistiendo para bien vivir*, significa la causa final perfecta de la misma, porque los que viven civilmente no sólo viven, lo que hacen las bestias y los esclavos, sino viven bien, se dedican, en efecto, a las actividades liberales, como son las de las facultades del alma, tanto de la práctica como de la especulativa.

§ 2. Así definida la ciudad por el vivir y el bien vivir como fin, conviene tratar primero del mismo vivir y de sus modos. Pues es aquello, como dijimos, por razón de lo cual se ha instituido la ciudad y es una necesidad que está en la base de todo lo que es y se hace por la comunicación de los hombres en ella. Y asentaremos como principio de todas las demostraciones, principio inserto en la naturaleza, creído y admitido por todos, que todos los hombres, no tarados ni impedidos por otra razón, desean naturalmente una vida suficiente y rehúyen y rechazan lo que la daña, lo cual no sólo se admite para los hombres, sino también para los animales de todo género, según Tulio, 1º *De los deberes*, cap. 3º, donde dice así: *Lo primero, proveyó la naturaleza a todo género de animales de lo necesario para que defiendan su cuerpo y su vida, y eviten aquello que les resulte nocivo, y adquieran y se proporcionen todo aquello que les es necesario para vivir*². Cosa que también, por inducción razonable, palmariamente puede cualquiera aceptar.

§ 3. Pero el vivir mismo y el bien vivir es conveniente al hombre de dos modos, uno temporal o intramundano, otro eterno, o, como se acostumbra a decir, celeste, y porque este segundo modo de vivir, a saber, el eterno, no lo pudieron persuadir por demostración la universalidad de los filósofos, ni es de las cosas manifiestas por sí mismas; por eso no se cuidaron de legarnos aquellas cosas que se hacen en fuerza de ese modo. Pero del vivir y del bien vivir o de la vida buena según el primer modo, el terrestre, y de las cosas que son necesarias para él, los filósofos ilustres tuvieron conocimiento por demostración de modo casi perfecto. Para conseguirlo llegaron a la conclusión de la necesidad de la comunidad civil, sin la cual esa suficiencia de vida no puede alcanzarse. De lo cual el eximio Aristóteles, 1º

¹ *Polit.*, l. 1, c. 2; 1252 b 27-29.

² *Cíc., De officiis*, I, 4.

de su *Política*, cap. 1° dijo: *Todos (los hombres) son llevados a ella a impulsos de la naturaleza con vistas a ese fin*³. De lo cual, aunque la experiencia razonable lo enseña, queremos con todo aducir la causa que dijimos, exponiéndolo más en detalle. Por nacer el hombre compuesto de elementos contrarios, por cuyas contrarias acciones y pasiones como que se corrompe continuamente algo de su substancia; y, además, por nacer desnudo e inerte, pasible y corruptible por el exceso del aire y de los otros elementos, como se dijo en la ciencia de las cosas naturales, necesitó de artes de diversos géneros y especies para defenderse de los daños dichos. Las cuales artes, no pudiendo ser ejercitadas sino por mucha gente, ni mantenerse sino por su recíproca comunicación, convino que los hombres se agruparan para tener la ventaja de esas cosas y apartar los inconvenientes.

§ 4. Mas como entre los hombres así congregados surgen contiendas y reyertas que, de no ser reguladas por las normas de la justicia, vendrían a generar luchas y disensiones entre los hombres y, finalmente, sería la ruina de la ciudad, convino establecer en esa comunidad una norma de lo justo y un guardián o ejecutor. Y como este guardián tiene el oficio de castigar a los delincuentes que se extralimitan y a los otros particulares que, dentro o fuera, intentan perturbar u oprimir a la comunidad, convino que la ciudad tuviese algo con qué resistir a éstos, y también, como la comunidad necesita algunos fondos de emergencia, repuestos y depósitos guardados de algunas cosas comunes, unas para el tiempo de paz, otras para el tiempo de guerra, fue necesario que hubiera en ella proveedores para poder acudir a la necesidad cuando fuere oportuno o necesario. Y fuera de lo dicho que mira a la sola necesidad de esta vida presente, hay otra cosa de la que necesitan los que conviven en su vida civil para el estado del siglo venidero, prometido al género humano por revelación sobrenatural de Dios y útil también para el estado de la vida presente, a saber, el culto y honor y acción de gracias rendidas a Dios, tanto por los beneficios recibidos en este mundo, como por los que se recibirán en el futuro; para enseñar esto y dirigir a los hombres convino que la ciudad designara algunos doctores. De todo esto y de lo demás apuntado antes, qué y cuáles sean esas cosas, puntualmente se dirá en los siguientes apartados.

§ 5. Hubo, pues, hombres asociados para tener una suficiencia de vida, con poder para procurarse las cosas necesarias antes reseñadas, comunicándoselas entre sí. Esta agrupación así perfecta y con una extensión suficiente se llamó ciudad, cuya causa final y plural diversidad de partes ya se ha declarado de algún modo y en lo siguiente se declarará más por

³ *Polit.*, l. 1, c. 2; 1253 a 29.

menudo. Porque siendo necesarias a los que quieren vivir suficientemente diversas cosas que no pueden procurarse por hombres de un solo orden u oficio, convino que hubiese diversos órdenes y oficios de hombres en esa comunidad, ejercitando y procurando esas tales cosas, de las que los hombres necesitan para su vida suficiente, y estos diversos órdenes u oficios no son otra cosa que la pluralidad y división de las partes componentes de la ciudad.

Qué sea, pues, la ciudad y por qué fue creada tal comunicación humana, y de la pluralidad y diferenciación de sus partes, baste haberlo así tocado de manera figurada.

V. DE LA DISTINCIÓN Y DEFINICIÓN DE LAS PARTES DE LA CIUDAD;
Y DE LA NECESIDAD DE SU EXISTENCIA Y SU DISTINCIÓN
POR EL FIN ASIGNADO POR LA INVENCION HUMANA

Una vez dicho ya en general lo que toca a las partes de la ciudad, en cuya acción y comunicación perfecta entre sí, con la exclusión de los impedimentos exteriores, dijimos que consistía su tranquilidad, para que quede más clara la causa de esa tranquilidad y de su opuesto, con una mayor determinación, tanto de parte de las obras o fines, como de las otras causas propias, volviendo al tema, digamos que las partes u oficios de la ciudad son de seis clases, como dijo Aristóteles en el 7º de la *Política*, cap. 6º: agricultores, artesanos, soldados, tesoreros, sacerdotes y jueces o consejeros. De las cuales clases tres, a saber, la sacerdotal, la militar y la judicial, son por excelencia partes de la ciudad, las que también, en las comunidades civiles, se dicen partes honorables¹. Las restantes se dicen partes en sentido lato, porque vienen a ser oficios necesarios en la ciudad, según Aristóteles en el 7º de la *Política*, cap. 7º Y a su conjunto suele denominársele vulgo. Son éstas las partes más conocidas de la ciudad o del reino y a ellas todas las otras pueden convenientemente reducirse. [...]

VI. DE LA CAUSA FINAL DE UNA PARTE DE LA CIUDAD, LA SACERDOTAL, FUNDADA
EN LA TRADICIÓN DIVINA O REVELACIÓN INMEDIATA, PERO IMPOSIBLE DE SER
PROBADA POR LA RAZÓN HUMANA

Resta de esta exposición decir de la causa final por la que fue instituido en la comunidad de los fieles el verdadero sacerdocio. Es la moderación de los actos humanos imperados por el conocimiento y el apetito, tanto

¹ *Polit.*, l. 7, c. 8; 1328 b 2; cfr. l. 4, c. 4; 1290 b 38.

de los inmanentes como de los transeúntes, según que por ellos el género humano se orienta al óptimo modo de vivir que es el del mundo futuro. Y por ello se ha de parar mientes en que, si bien el primer hombre, Adán, fue principalmente creado para la gloria de Dios, como las demás criaturas, fue, sin embargo, hecho de un modo singular distinto de las otras especies corruptibles, pues fue hecho a imagen y semejanza de Dios; capaz así y partícipe de la felicidad eterna después de la vida en este mundo. Fue también creado en estado de inocencia y justicia original y también de gracia, como con probabilidad afirman algunos de los santos y algunos insignes doctores de la sagrada Escritura. En el cual estado, si hubiera continuado, no le hubiera sido necesaria a él ni a su posteridad la institución y distinción de los oficios civiles, por razón de que la naturaleza hubiera producido para él todo lo necesario y deleitoso para una suficiencia de vida en el paraíso terrenal o de deleites, y eso sin ninguna pena ni fatiga suya.

§ 2. Pero como corrompió su inocencia o su original justicia y gracia al comer del árbol prohibido, faltando con ello al mandato divino, al punto cayó en la culpa y en la miseria o pena, quiero decir, la pena de la privación de la felicidad eterna a la cual él con su posterioridad había sido ordenado como a su fin último por la bondad del Dios de la gloria. De esta trasgresión del mandato dicho se mereció el propagar por medio de la concupiscencia toda su posterioridad, por la cual y en la cual concupiscencia todo hombre después es engendrado y nacido, contrayendo por esto el pecado que en la ley de los cristianos llamamos original, excluido Cristo Jesús, que fue concebido sin pecado ni concupiscencia alguna por obra del Espíritu Santo y nacido de María Virgen; lo cual acaeció cuando una de las tres divinas personas, a saber, el Hijo, verdadero Dios, asumió en la unidad de su persona la naturaleza humana. Pues de esta trasgresión de los primeros padres quedó enferma en el alma toda la posteridad humana y enferma nace la que fuera creada en estado de salud perfecta de inocencia y de gracia, privada, asimismo, por el pecado, del último felicísimo fin al que había sido ordenada.

§ 3. Siendo propio de Dios compadecerse del género humano, hechura e imagen suya, la que había destinado él a la vida bienaventurada y eterna, el que nunca hace nada en vano ni falla en lo necesario, quiso poner remedio en el caso del hombre, dando ciertos preceptos de obediencia que guardara el hombre, los que, contrarios a la trasgresión, vinieran a sanar la enfermedad de la culpa que de aquélla provenía. Y procedió muy ordenadamente, como diestro médico, desde lo más fácil a lo más difícil. Primero prescribió a los hombres el rito de los holocaustos, de las primicias de los

frutos de la tierra y de los primogénitos de los animales, como queriendo poner a prueba la penitencia y la obediencia humana. Este rito lo guardaron los padres antiguos en reconocimiento, fe, obediencia y acción de gracias a Dios, hasta los tiempos de Abraham. A éste, según lo dicho, le dio un precepto más grave, a saber, la circuncisión de todo sexo masculino humano en la carne del prepucio, como queriendo Dios probar de nuevo más ampliamente la penitencia y obediencia humana. Fueron estos preceptos observados por algunos hasta los tiempos de Moisés, por medio del cual dio Dios después al pueblo de Israel una ley, en la que, además de los dichos, puso más amplios preceptos, tanto para la vida presente como para la futura, y de la misma ley estableció también, como ministros, a sacerdotes y levitas. La utilidad de estos primeros preceptos, y de los de la ley mosaica, consistía en una cierta purificación del pecado o culpa, tanto de la original como de la actual o cometida voluntariamente, y la fuga o preservación de la pena eterna y temporal del otro mundo, aunque por estas observancias no merecieran los hombres la eterna felicidad.

§ 4. Pero una vez que el misericordioso Dios había destinado el género humano a esa felicidad, queriendo volverle de la caída y restituirle a la misma según un orden conveniente, al final de todo, por medio de su Hijo Jesucristo, verdadero Dios y verdadero hombre en la unidad de la persona, dio la ley evangélica, que contiene preceptos respecto de lo que se ha de creer, hacer y rechazar, y juntamente consejos. Con cuyo cumplimiento, no sólo se libran los hombres de la pena, como por la anterior observancia, sino que, por una graciosa disposición, merecen por ello, por cierta congruencia, la felicidad eterna. Y por ello se la llama ley de gracia, ya porque por la pasión y muerte de Cristo fue redimido el género humano de la culpa y pena de daño que priva de la eterna felicidad, en la cual había incurrido por la caída y pecado de los primeros padres, ya porque mediante su observancia y la recepción de los sacramentos instituidos en ella y por ella se nos confiere la gracia divina, dada se corrobora, y perdida se recupera, y por la cual, por disposición de Dios con el mérito de la pasión de Cristo, por cierta congruencia, nuestras obras, como dijimos, se hacen merecedoras de la felicidad eterna.

§ 5. Por el mérito de esta pasión de Cristo, no sólo los posteriores recibieron la gracia, con la que pueden merecer la vida bienaventurada, sino también los cumplidores de los mandatos de la ley de Moisés consiguieron la gracia de la eterna bienaventuranza, de la cual permanecían privados en el otro mundo, en el lugar que llaman limbo, hasta el advenimiento, pasión, muerte y resurrección de Cristo. Por él recibieron la promesa dada por Dios, aunque en los preceptos primeros de los patriarcas y de la ley de

Moisés la promesa de tal gracia se les daba bajo un velo de enigma, porque *todo les ocurría en figura*, como dice el apóstol a los Romanos [Corintios]¹.

§ 6. Y fue muy conveniente este plan restaurador divino, porque de lo menos perfecto fue a lo más perfecto y, finalmente, a lo perfectísimo de lo conducente a la humana salvación. Ni se ha de opinar que no pudiera Dios, si quisiera, en seguida, al principio, tras la caída del hombre, poner remedio. Sino que lo hizo así porque así lo quiso y porque convino, pues exigía el pecado de los hombres que la facilidad del perdón demasiado pronto no diera ocasión para delinquir en lo sucesivo.

§ 7. De la dicha ley se instituyeron doctores y, de acuerdo con ella, administradores de los sacramentos en las comunidades, llamados sacerdotes, diáconos o levitas, cuyo oficio es enseñar los preceptos y los consejos de la ley evangélica cristiana, en aquellas cosas que hay que creer, que hacer o que evitar, teniendo como fin el conseguir el estado feliz del mundo futuro y evitar lo opuesto.

§ 8. El fin, pues, sacerdotal es la enseñanza y la información de lo que, según la ley evangélica, es necesario creer, hacer y omitir para conseguir la eterna salvación y huir de la perdición.

[...]

VII. DE LAS OTRAS CLASES DE CAUSAS DE LA EXISTENCIA Y DISTINCIÓN DE LAS PARTES DE LA CIUDAD Y DE LA DIVISIÓN, DENTRO DE CADA UNA DE ESAS PARTES, SEGÚN DOS MODOS QUE HACEN A NUESTRO PROPÓSITO

Consiguientemente a esto toca ahora hablar de las otras causas de los oficios o partes de la ciudad, y primero hablaremos de las causas materiales y formales, luego de la motora o eficiente de las mismas. Y como en las cosas que tienen su realización mediante la mente humana la materia preexiste a la forma en acto, hablaremos primero de la causa material. Y diremos que la materia propia de los diversos oficios, en la medida en que los oficios se denominan hábitos del alma, son los hombres, inclinados por su misma generación o nacimiento a diversas artes o disciplinas. Porque, no faltando la naturaleza en las cosas necesarias, y siendo más cuidadosa en las más nobles, como entre los seres corruptibles es la especie humana, de la cual, perfeccionada por las diversas artes y disciplinas, es conveniente que, como de materia, se constituya la ciudad y sus partes en ella necesarias para obtener la suficiencia de vida, como en los caps. IV y V se ha mostra-

¹ I Cor, 10, 11.

do, inició ella misma, por la generación de los hombres, esta distinción y división, produciendo, según naturales disposiciones, unos aptos e inclinados a la agricultura, otros a la milicia, otros a diversos géneros de artes y disciplinas, diversidad de hombres para la diversidad de funciones. Y no uno solo a determinado arte o disciplina de una sola clase, sino muchos a la misma clase de arte o disciplina, según la necesidad de la suficiencia lo requería. A algunos, así pues, engendró aptos para la prudencia, porque de prudentes debe componerse en la ciudad la parte judicial y deliberativa; a otros los engendró aptos para el vigor y la audacia, pues de los tales convenientemente se constituye la parte militar. A los demás los adaptó para las clases de hábitos prácticos y especulativos necesarios o convenientes para el vivir y para el bien vivir; de forma que, contando con la diversidad de inclinaciones naturales a los diversos géneros y especies de hábitos, se cumpliese en todo lo necesario a la diversidad de las partes de la ciudad. Y así aparecen ya bastante claras las causas materiales de los oficios de la ciudad, según que suelen llamar oficios a las partes de la ciudad. Estas son los hombres experimentados en las artes y disciplinas de diversos géneros y especies, de los que se forman los diferentes órdenes o partes en la ciudad, con miras a las suficiencias finales provenientes de las artes y disciplinas de ellos; por donde con propiedad se dicen partes de la ciudad los oficios, como si dijéramos servicios, pues mirados como están establecidos en la ciudad, se ordenan al servicio y regalo de los hombres.

§ 2. Las causas formales, en cuanto hábitos de la mente humana, no son otros que estos mismos hábitos; pues los tales son, en los que los poseen, formas complementarias o perfectivas de las inclinaciones humanas que existen por naturaleza. De donde en la *Política*, 7^o, cap. final: *Toda arte y disciplina trata de suplir lo que falta a la naturaleza*¹. Las causas formales de los oficios, en cuanto han sido instituidos y son partes de la ciudad, son las normas directivas de la causa eficiente, transmitidas o impresas en aquellos que en la ciudad están designados para llevar a cabo determinadas obras.

§ 3. Las causas motoras o ejecutoras de los oficios, en cuanto significan hábitos del alma, son las voluntades de los hombres, sus conocimientos y deseos, o cada uno de por sí, o todos juntos indistintamente. De algunos es también principio el movimiento y el ejercicio de los órganos corporales. Pero su causa eficiente, en cuanto son partes de la ciudad, es el legislador humano frecuentemente y las más de las veces, si bien en otro tiempo y rara vez y en poquísimas cosas de alguno o de algunos fue Dios la

¹ *Polit.*, l. 7. c. 17; 1337 a l.

causa motora inmediata, sin humana determinación alguna, como se dirá en el XII y XVII de la Segunda Parte. Sobre el sacerdocio hay otra forma de institución y de ella se dirá ampliamente en el XV y XVII de la Segunda Parte.

VIII. DE LOS GÉNEROS DE GOBIERNO O RÉGIMEN POLÍTICO;
TEMPLADO Y VICIADO, Y DIVISIÓN DE SUS CLASES

En lo que antecede hemos ya mostrado de algún modo, pero es preciso mostrar con más evidencia, que la institución y diversificación de las partes de la ciudad se realiza por la acción de alguna causa motora, a la que hemos llamado antes legislador. Y como el mismo legislador establece, distingue y separa estas partes a la manera de la naturaleza de un viviente, formando primero y estableciendo en la ciudad una parte, que en el cap. V de esta Parte llamamos gobernante o judicial, y por ella las demás, como se declarará más en el XV de esta Parte, conviene que digamos primero algo de la naturaleza de esta parte. Pues siendo la primera de todas, como aparecerá de lo que sigue, de la declaración primero de su eficiencia, convenientemente entraremos en la declaración de la institución y distinción activa de las otras partes de la ciudad.

§ 2. Hay, pues, dos géneros de la parte gobernante o gobierno, uno templado y otro viciado; llamo *bien templado* a uno de ellos con Aristóteles en el 3º de la *Política*, cap. 5º, a aquél en el que el príncipe gobierna mirando a lo útil a la comunidad según la voluntad de los súbditos, y *viciado* llamo al que falla en esto¹. Cada uno de estos géneros se divide en tres especies; el primero, el templado, en monarquía regia, aristocracia y república; el otro, el viciado, en otras tres clases opuestas, tiranía monárquica, oligarquía y democracia. Y cada una de estas clases tiene sus modos, sobre los que por el momento no es cuestión de hablar por menudo. De ellos trató abundantemente Aristóteles en el 3º y 4º de su *Política*².

§ 3. A fin de tener más pleno conocimiento de esas clases, lo que será necesario de algún modo para lo que a seguido hemos de declarar, describimos cada una de esas formas de gobierno según la intención de Aristóteles, diciendo, lo primero, que la monarquía regia es un modo templado de gobierno en el que uno solo manda para el común provecho, con la voluntad y consenso de los súbditos. La tiranía a él opuesta es un gobierno viciado en el que uno solo manda para el provecho propio sin contar con la voluntad de los súbditos. La aristocracia es un gobierno templado, en el

¹ *Polit.*, 1. 3, c. 6; 1279 a 17-22.

² *Polit.*, 1. 3, c. 7-9; 1279 a 22 – 1279 b 10; 1285 b 35.

que manda una sola clase honorable de acuerdo con la voluntad de los súbditos, o según el consenso y el provecho común. La oligarquía a ella opuesta es un gobierno viciado en el que mandan algunos de entre los más ricos o más poderosos, mirando al provecho de ellos, sin contar con la voluntad de los súbditos. La república, aunque en una acepción del vocablo designa algo común a todos los géneros o formas de gobierno o régimen, contraída a una especial significación, importa un modo de gobierno templado en el que todo ciudadano participa de algún modo en el gobierno o en el poder consultivo, según el grado, haberes y condición del mismo, mirando al común bien y de acuerdo con la voluntad y consenso de los ciudadanos. La democracia, a ella opuesta, es el gobierno en el que el vulgo, o la multitud de pobres, impone su gobierno y rige sola sin contar con la voluntad y consenso de los demás ciudadanos, ni absolutamente mira al común bien según una justa proporción.

§ 4. Cuál sea, entre los templados el mejor gobierno, o cuál, entre los viciados el peor, y del grado de bondad o malicia de los otros, no toca decirlo en esta disertación. Baste lo dicho sobre la división de los regímenes de gobierno según sus formas y sobre la descripción de ellas.

IX. DE LOS MODOS DE INSTITUIR LA MONARQUÍA; REGIA Y DEFINICIÓN DE SU PERFECCIÓN. MODOS TAMBIÉN DE INSTITUIR LOS OTROS RÉGIMENES O FORMAS DE GOBIERNO, LAS TEMPLADAS Y LAS VICIADAS

Dejando esto asentado, consiguientemente se ha de decir de los modos de crear o instituir la parte gobernante. Porque procediendo de su naturaleza, mejor o peor, lo que de allí sale, como las acciones para el régimen civil, conviene deducir la causa agente, de la que tanto esos modos como la parte gobernante, por medio de ellos, han de provenir para la mayor utilidad de la cosa pública.

§ 2. Pero como lo que pretendemos tratar en este libro es de las causas y acciones por las que lo más corrientemente debe ser creada esa parte gobernante, queremos primero hablar del modo y causa por la cual fue ya creada esa parte gobernante, aunque rara vez, a fin de que diferencemos ese modo y esa acción y su causa inmediata, de los modos y acciones y sus causas inmediatas por las cuales regularmente y por la mayor parte se debe instituir, y las que con humana demostración podemos probar. Porque del primer modo no se puede tener un conocimiento cierto por demostración humana. Pues este modo o acción y su causa inmediata, por la que fue ya creada la parte gobernante y las otras partes de la ciudad, principalmente

el sacerdocio, fue la voluntad divina mandándolo por el oráculo determinado de alguna singular criatura, o acaso por sí misma inmediatamente; del cual modo instituyó el gobierno del pueblo israelítico en la persona de Moisés y de algunos otros jueces después de él y el sacerdocio en la persona de Aarón y de sus sucesores. De la causa y su acción libre, mostrar o decir por qué obró así o de otro modo y por qué no se hace o se hizo así, no podemos decir nada por demostración, sino con simple fe y sin apelar a la razón lo admitimos. Otra es la institución de los regímenes que inmediatamente proviene de la mente humana, aunque venga de Dios como de causa remota, quien otorga también todo principado terreno, como en Juan, 19°¹, y claramente lo dice el apóstol a los Romanos, 13°² y el bienaventurado Agustín en el 5° De la ciudad de Dios, cap. 21°³, lo que no se hace siempre inmediatamente, sino las más de las veces y en todas partes lo instituyó Dios por medio de las mentes de los hombres, a los que confió el arbitrio de tal institución. Y de esta causa, cuál sea y con qué género de acción deba instituir tales cosas, reparando en lo mejor o en lo peor para la realidad política, puede ser determinado por demostración por la humana certeza.

§ 3. Omitiendo, pues, el modo que no podemos evidenciar por demostración, vamos a enumerar, primero, los modos de la institución de los regímenes creados por la voluntad humana; luego mostraremos cuál de ellos es más cierto y simple; después, partiendo de la naturaleza mejor de ese modo, argüiremos la causa motora de la que únicamente debe llevar la institución mejor y a la determinación de las otras partes de la ciudad. Finalmente, diremos lo tocante a la unidad del gobierno, de la cual también aparecerá cuál es la unidad de la ciudad o del reino.

§ 4. Siguiendo, pues, el propósito, enumeraremos, lo primero, los modos de la institución de la monarquía regia, hablando de su origen. Porque esta clase de gobierno aparece al pronto como connatural a nosotros y próxima a la organización doméstica, como consta de lo dicho en el cap. III. Después, de esta determinación aparecerá consiguientemente lo tocante a los modos de la institución de las otras divisiones del gobierno. Son, pues, los modos o las instituciones posibles de la monarquía regia cinco en número, según Aristóteles, 3° de la *Política*, cap. 8°⁴. Uno, cuando la monarquía se establece para una determinada empresa, pero que afecta al régimen de la comunidad, como la conducción del ejército, bien con sucesión heredita-

¹ Juan, 19, 1.

² Rom., 13, 1.

³ De civ. Dei, 5, 21.

⁴ Polit., l. 3, c. 14; 1284 b 35.

ria, o bien para el tiempo de una sola persona, como se instituía a Agamenón jefe del ejército por los griegos. Se designa este oficio en las comunidades modernas con los nombres de capitán o de condestable. Y este jefe del ejército, luego, en tiempo de paz, no se entrometía en ningún juicio público, pero durante el tiempo que militaba en el ejército era dueño de matar y de castigar de otras maneras a los trasgresores.

Otro modo es cómo gobiernan las monarquías de Asia, que tienen por herencia de sus predecesores el dominio, con arreglo siempre a una ley adaptada al monarca más que absolutamente a la comunidad, monarquía casi despótica. Porque los habitantes de aquella región toleran semejante gobierno sin pesar suyo por su bárbara y servil condición y ayuda a ello la costumbre. Y éste es principado regio, porque es heredado y cuenta con súbditos voluntarios, puesto que los primeros habitantes de la región fueron antecesores del monarca; pero es en algún modo tiránico por no estar sus leyes encaminadas absolutamente al bien común, sino al del monarca.

Un tercer modo de gobierno regio es cuando gobierna alguien elegido, no por herencia patria o paterna, pero según una ley cuasi tiránica, por no estar ordenada absolutamente al bien común, sino más al del monarca. Por lo cual llamó Aristóteles en el mismo lugar a esta clase de gobierno *tiranía electiva*, tiranía por el carácter despótico de la ley, electiva porque no es contra la voluntad de los súbditos.

El cuarto modo es aquél en que por elección se instituye a alguno con derecho a sucesión para toda su prole, según una ley que mira absolutamente al bien común. Y éste era el usado en los tiempos heroicos, como dice Aristóteles en el mismo lugar. Se llamaron tiempos heroicos aquéllos, o porque la constelación producía tales hombres que se creía que eran héroes, es decir, divinos, porque los tales, y no otros, eran constituidos príncipes por lo extraordinario de su virtud y sus acciones benéficas, como reunir una multitud dispersa y agruparla en comunidad civil, o porque por la lucha y valentía de sus armas libraron la región de los opresores, o porque quizá compraron el terreno o lo adquirieron de otro modo conveniente y lo distribuyeron entre los súbditos. y para decirlo de una vez, por la prestación de un gran beneficio, o sobresalir en alguna otra virtud en favor de la restante multitud, fueron éstos creados gobernantes, con sucesión de sus herederos todos, como también dice Aristóteles en el 5° de la *Política*, cap. 5°⁵. Bajo esta forma de monarquía también comprende quizá Aristóteles aquélla en la que alguien es elegido por el tiempo de su vida o por un período de tiempo, y quizá Aristóteles nos la quiso dar a entender refiriéndonos a ésta y a la llamada tiranía electiva, por participar de ambas.

⁵ *Polit.*, I, 5, c. 10; 1310 b 10.

El quinto modo es y fue aquél en el que el gobernante es constituido señor de todo lo que hay en la comunidad, disponiéndolo todo a su voluntad, las cosas y las personas, como un jefe de familia que dispone a su arbitrio de todo lo que hay en su casa.

§ 5. Para poner más en claro lo dicho por Aristóteles y para reducir a un capítulo los modos todos de instituir los otros sistemas de gobierno, diremos que todo gobierno, o es conforme a la voluntad de los súbditos, o es sin su voluntad. El primero es el género de los gobiernos bien temperados, el segundo el de los viciosos. Y cada uno de esos géneros se divide en tres especies o modos, como se dijo en el cap. VIII. Y como una especie de forma de gobierno bien templada y quizá la más perfecta es la monarquía regia, por ello, resumiendo lo dicho antes y partiendo de sus modos, comencemos diciendo que el rey o monarca, o es constituido por elección de los habitantes o ciudadanos, o sin esa elección obtuvo el gobierno de modo justo. Si sin la elección de los ciudadanos, o es así porque habitó él primero la región o sus antecesores de los que trae el origen, o porque compró la tierra y el derecho sobre ella, o la adquirió en justa guerra, o de otro modo justo, por ejemplo por donación a él hecha por servicio prestado. Participa cada uno de estos modos tanto más del verdadero modo de monarquía regia cuanto más se relaciona con sus súbditos voluntarios y según ley dirigida al provecho común de los súbditos; y tanto más sabe a tiranía cuanto más se aparta de eso, a saber, del consentimiento de los súbditos y de la ley instituida para el provecho de los mismos. De donde en el 4° de la *Política*, cap. 8° se escribe: *Eran regias porque gobernaban según la ley*, las monarquías, se entiende, y *porque monarquizaban a súbditos voluntarios, y tiránicas, porque gobernaban despóticamente y según su arbitrio*⁶, de los monarcas, se entiende. Estas dos cosas dichas dividen el gobierno templado del viciado, como aparece de la clara sentencia de Aristóteles, según que se dé o totalmente o en mayor grado el consentimiento de los súbditos.

Si el monarca reinante fue instituido por la elección de los habitantes, o se constituye con derecho de sucesión para toda su descendencia, o no. Si no se constituye con derecho de sucesión para toda su descendencia, esto puede ocurrir de dos modos, o porque se constituye para sólo el tiempo de vida de uno, o de uno o algún sucesor suyo, uno o muchos, o no se constituye para toda la vida de uno o alguno de sus sucesores, ni del primero ni de los siguientes, sino solamente para un período de tiempo determinado, como un año o dos, o más o menos, y todavía, o para juzgar y resolver en toda clase de asuntos, o para uno solo, como capitanear el ejército.

⁶ *Polit.*, 1. 4: 10; 1295 a 15.

§ 6. Coinciden y difieren estos gobiernos monárquicos electivos y no electivos en que unos y otros gobiernan a súbditos voluntarios, y difieren en que, como ocurre las más de las veces, en los no electivos gobiernan a súbditos menos voluntarios y los rigen con leyes menos políticas, atentas al bien común, como dijimos antes sobre las naciones bárbaras. Los electivos gobiernan con más conformidad de los súbditos y los rigen con leyes más políticas, atentas al bien común.

§ 7. De lo que aparece claro, lo que se declarará más en lo que sigue, que el modo de gobierno electivo aventaja al no electivo. Y es sentencia de Aristóteles en el 3º de la *Política*, cap. 8º, que adujimos más arriba sobre las formas de gobierno de las edades heroicas⁷. Además este modo de institución es más estable en las comunidades perfectas. Pues todos los otros habrá que reducirlos de necesidad a éste alguna vez, no al revés, como si falta la sucesión del heredero o por otra causa se hace aquel sistema insoportable a la multitud por sobra de perversidad de su gobierno; conviene entonces que la multitud vuelva los ojos a la elección, la cual nunca puede faltar, no agotándose la generación de los hombres. Además sólo de este modo de institución se obtiene el mejor gobernante. Ha de ser, en efecto, el mejor de aquellos que se ocupan de la cosa pública, pues ha de regir los actos civiles de todos los otros.

§ 8. El modo de institución de las otras formas de gobierno templado, por la mayor parte es la elección y alguna vez la suerte, sin sucesión familiar hereditaria. Los modos de institución de las otras formas de gobierno viciado, muy frecuentemente son el fraude o la violencia, o ambas cosas.

§ 9. Cuál de las formas de gobierno templado sea la mejor, la monarquía, o las otras dos, la aristocracia o la república, y todavía, entre las monarquías, si la electiva o la no electiva, y entre las electivas, si la que se instituye con plena sucesión hereditaria, o aquella en que uno solo sin tal sucesión, la cual se subdivide aún en dos, la que es para toda la vida de alguno o algunos, y la que es para un período, anual, bienal, o más largo o más corto, es cuestión de considerarlo, y no carece de duda razonable, aunque sin dudar ha de tenerse como verdad y conforme a la sentencia de Aristóteles, que la elección es la más cierta regla de cualquier forma de gobierno, como en el XII, XVI y XVII de esta Parte se pondrá más de manifiesto.

§ 10. Pero no hemos de ignorar que otras y otras gentes, en diversas regiones y tiempos, están más dispuestas a otras y diversas clases de política y a otras y otras formas de gobierno, como dice Aristóteles en el 3º de la

⁷ *Polit.*, l. 3, c. 14; 1285 b 2.

Política, cap. 9º, lo que ha de tenerse en cuenta por los legisladores y los creadores del régimen político⁸. Porque como no todo hombre está dispuesto para la mejor disciplina, y por ello no es orientado a ella con facilidad por el que lo dirige, sino a aquélla para la que, entre las buenas, esté mejor preparado, así acaso una multitud alguna vez o en algún lugar no está dispuesta a tolerar la mejor forma de gobierno, y por ello habrá que llevarla primero a experimentar la más adecuada a ella de entre las templadas. Porque antes de la monarquía de Julio César el pueblo romano no toleró mucho tiempo a un monarca determinado, ni hereditario, ni para todo el tiempo de vida de uno solo. Lo cual quizá le aconteció por la multitud de los hombres heroicos y genios políticos, ya como familias o parentescos, ya como personas individuales.

§ 11. De lo que hemos dejado sentado aparece claro que los que se preguntan cuál de los monarcas sea mejor para la ciudad o el reino, el electivo o el hereditario, no llevan buen camino en su pregunta. Sino que conviene que, corrigiendo, se pregunten primero qué monarca es mejor, el electivo o el no electivo. Y si el electivo, todavía cuál de los elegidos, si el que se establece con sucesión hereditaria, o sin ella. Porque si el monarca no elegido casi siempre transmite el poder al heredero, el elegido no siempre, sino sólo cuando se le constituye en gobierno con pleno derecho de sucesión.

Sobre los modos de constitución de los gobiernos y de que aventaja entre ellos absolutamente el de elección, quede así dicho.

X. DE LA DISTINCIÓN Y DEFINICIÓN DE LOS SIGNIFICADOS DE LA PALABRA LEY Y SU PROPÍSIMA ACEPCIÓN CONFIRMADA POR NOSOTROS

Tal como afirmamos que la elección es el más perfecto y ventajoso de los modos de establecer el gobierno, bien será inquirir en su causa eficiente, de la cual efectivamente provenga la realización plena de su bondad. Porque de aquí aparecerá también la causa del gobierno elegido y similarmente la de las otras partes de la ciudad. Pero porque el gobierno ha de regular los actos humanos civiles, como demostramos en el V de esta Parte, y esto, según la regla que ha de ser forma del gobernante en cuanto tal, hay que investigar esta regla, si existe, cuál es y para qué es. Porque acaso sea la misma causa eficiente, la del gobierno y la del gobernante.

§ 2. Suponiendo, pues, como cosa evidente por inducción de todas las comunidades perfectas, que existe esa regla, a la que llaman estatuto o

⁸ *Polit.*, l. 3, c. 14; 1284 b 37.

costumbre, o, con nombre común, ley, mostraremos, primero qué sea, luego su necesidad de fin, y por último concluiremos, por demostración, con qué clase de acción, por quién o por quiénes haya de ser establecida. Lo cual equivaldrá a investigar sobre el legislador o la causa eficiente del mismo, a la cual pensamos que pertenece también la elección de los principados, y lo mostraremos por demostración en lo que sigue. De lo que también aparecerá la materia o el sujeto de dicha regla que denominamos ley. Pues ella es la parte gobernante, a la que pertenece regular, según ley, los actos políticos y civiles de los hombres.

§ 3. Procediendo, pues, a lo dicho, para que no surja confusión por la multiplicidad del nombre, conviene distinguir las intenciones o significaciones de este nombre, *ley*. Porque este nombre, entre las múltiples acepciones, importa, en uno de sus significados, la natural inclinación a alguna acción o pasión, como la llamó el apóstol a los Romanos, 7°, cuando dijo: *Pero veo otra ley en mis miembros que contradice a la ley del espíritu*¹. Con otra acepción se dice *ley* de cualquier hábito operativo y, generalmente, de toda forma de cosa factible, existente o en la mente, de la que provienen, como de su ejemplar, la medida de la forma de los artefactos, como dice Ezequiel, 4°: *Esta es la ley de la casa, éstas las medidas del altar*². En tercer lugar se toma *ley* por regla que contiene las normas de los actos imperados humanos, según que son ordenados a la gloria o a la pena en el mundo venidero, a tenor de la cual significación la ley mosaica se dijo ley, en cuanto a una parte de ella, mientras la ley evangélica toda ella se dice ley. De donde el apóstol, hablando de ellas a los hebreos, dice: *Trasferido el sacerdocio, es necesario que se transfiera también la ley*³. Y así también en la enseñanza evangélica se dice ley en Santiago, 1°: *Y el que mirare a esta ley perfecta de la libertad y en ella permaneciere, etc., será bienaventurado en su acción*⁴. Según esta acepción de la ley, las sectas o religiones se dicen leyes, como la de Mahoma o la de los persas, todas o algunas de sus partes, aunque de éstas sólo la mosaica y la evangélica, es decir, la cristiana, contiene la verdad. Y así también llamó Aristóteles leyes a las sectas religiosas en el 2° de la *Filosofía [Met.]*, cuando dijo: *Cuánta fuerza tenga lo que se hace costumbre lo muestran las leyes*⁵ y en el 12° de la misma: *Todo lo demás fue introducido míticamente para persuadir a la multitud en orden a las leyes y a lo útil común*⁶. En cuarto lugar importa

¹ Rom., 7, 23.

² Ez., 43, 12-13.

³ Hebr., 8-12.

⁴ Sant., 1-25.

⁵ Met., l. 2, c. 3; 995 a 4.

⁶ Met., l. 12, c. 8; 1074 b 3.

este nombre *ley*, como lo más notorio, la ciencia o la doctrina o el juicio universal de lo justo y civilmente útil y de sus opuestos.

§ 4. Y tomada así la ley, puede considerarse en dos maneras, una en sí misma, en cuanto por ella solamente se muestra lo que es justo o injusto, útil o nocivo, y como tal se dice ciencia o doctrina del derecho. La otra manera de considerarla es cuando para su observancia se da un precepto coactivo con pena o premio en este mundo, o en cuanto se da en forma de tal precepto, y de este modo considerada se dice y es propísimamente ley. A ésta así tomada la define Aristóteles en el último de la *Ética*, cap. 8º, cuando dice: *La ley contiene una fuerza coactiva, siendo un enunciado emanado de una cierta prudencia e inteligencia; un enunciado, pues, o proposición procedente de una prudencia e inteligencia*, política, se entiende; o también, una ordenación sobre lo justo y lo conveniente y sus opuestos, según la prudencia política, *provista de fuerza coactiva*, es decir, sobre cuya observancia se da un precepto que cada cual ha de cumplir, o dada por modo de tal precepto, es la ley⁷.

§ 5. De donde no todos los conocimientos verdaderos de lo justo y lo conveniente civil son leyes, si no hay de ello un precepto coactivo, o no se dieran por modo de precepto, si bien ese conocimiento verdadero necesariamente se requiere para la ley perfecta. Más aún, a veces se dan por leyes falsas ideas de lo justo y lo útil cuando de ellas se da precepto de observancia, o se dictan por modo de precepto, como aparece en los países de algunos bárbaros que hacen observar como justo el que se absuelva de culpa y pena civil al homicida que ofrece un rescate por ese delito, cuando eso es enteramente injusto y por consiguiente sus leyes no son en absoluto perfectas. Aun cuando tengan la forma debida, a saber, un precepto coactivo obligante a su observancia, carecen, sin embargo, de la requerida condición, a saber, de la debida y verdadera ordenación de lo justo.

§ 6. Bajo esta acepción de la ley se comprenden todas las reglas de lo justo y útil civil, instituidas por la autoridad humana, como las costumbres, los estatutos, los plebiscitos, las decretales y todas las semejantes que, según hemos dicho, se apoyan en la autoridad humana.

§ 7. No debemos, con todo, ignorar que tanto la ley evangélica como la mosaica y quizá las demás de las religiones, diversamente consideradas y referidas en todo o en parte a los actos humanos en este mundo o en el venidero, unas veces vienen o vinieron o vendrán a colocarse en la tercera significación de la ley y otras en la última como se declarará más en el VIII y IX de la Segunda Parte; unas, en efecto, serán verdaderas, otras serán cosa de falaz imaginación y vanas promesas. Que exista, pues, alguna regla

⁷ *Eth. Nic.*, l. 10, c. 9; 1180 a 21.

o ley de los actos humanos civiles y en qué consista, por lo dicho queda claro.

XI. DE LA NECESIDAD DE DAR LEYES TOMADAS EN SENTIDO PROPÍSIMO,
Y QUE NINGÚN GOBERNANTE, POR VIRTUOSO Y JUSTO QUE SEA,
DEBE GOBERNAR SIN LEYES

[...]

§ 5. Queda mostrar que el gobernar según la ley y no fuera de ella, es deber de todos los gobernantes y muchísimo más de aquellos monarcas que gobiernan con sucesión hereditaria a fin de que su gobierno sea más seguro y más duradero, lo que se había afirmado en el comienzo de este capítulo como la segunda razón de la necesidad de las leyes. Y esto se puede ver, lo primero, porque gobernar según la ley preserva sus juicios de los defectos que ocurren por la ignorancia o por la pasión desordenada. De donde, regulados en sí mismos y en orden a sus súbditos, menos padecerán las sediciones y, consiguientemente, la disolución de los gobiernos que acontecerían obrando torcidamente según su arbitrio, como palmariamente lo dice Aristóteles en el 5° de la *Política*, cap. 5°: *Porque el reino, dice Aristóteles, sólo muy raramente se corrompe desde el exterior, pero desde dentro pueden venir muchas corrupciones. y se corrompe de dos modos; uno, por las sediciones de los que participan del poder regio, otra, por los que intentan gobernar más tiránicamente cuando se empeñan en ocuparse absorbentemente de demasiados asuntos y al margen de la ley. No se crean hoy muchos reinos, pero si se crean, son monarquías o más bien tiranías*¹.

[...]

XII. DE LA CAUSA EFICIENTE DEMOSTRABLE DE LAS LEYES HUMANAS
Y DE AQUELLA QUE NO PUEDE EVIDENCIARSE POR DEMOSTRACIÓN;
LO QUE EQUIVALE A INQUIRIR SOBRE EL LEGISLADOR.

DE DONDE APARECE TAMBIÉN QUE POR LA SOLA ELECCIÓN, SIN OTRA
CONFIRMACIÓN, SE DA LA AUTORIDAD A AQUEL QUE SE ESTABLECE POR ELECCIÓN

[...]

§ 2. Pues entrando en ello, digamos que descubrir esta ley tomada materialmente y según la tercera significación, a saber, como ciencia de lo justo y lo útil civil, es competencia de cualquier ciudadano, aunque más conveniente y adecuadamente puede hacerse partiendo de la observación de

¹ *Polit.*, 1. 5, c. 10; 1312 b 38 – 1313 a 5.

los que tienen posibilidad de vacar a ello, de los ancianos y experimentados en las cosas prácticas, los llamados prudentes, más que de la consideración de los de oficios mecánicos, los que se aplican a procurar con su trabajo las cosas necesarias para la vida. Pero como el conocimiento y el descubrimiento verdadero de lo justo y lo útil y de sus opuestos no es ley según su última y propia significación, con la que se constituye en medida de los actos humanos civiles, mientras no se dé un precepto coactivo de su observancia, o por modo de tal precepto se promulgue por aquél con cuya autoridad deben y pueden ser castigados los trasgresores, por eso conviene decir de quién o de quiénes sea la autoridad de dar tal precepto y de castigar sus trasgresores. Lo que equivale indagar el legislador o autor de la ley.

§ 3. Digamos, pues, mirando a la verdad y al consejo de Aristóteles en el 3º de la *Política*, cap. 6º¹ que el legislador o la causa eficiente primera y propia de la ley es el pueblo, o sea, la totalidad de los ciudadanos, o la parte prevalente de él, por su elección y voluntad expresada de palabra en la asamblea general de los ciudadanos, imponiendo o determinando algo que hacer u omitir acerca de los actos humanos civiles bajo pena o castigo temporal; digo la parte prevalente, atendida la cantidad y la calidad de las personas en aquella comunidad, para la cual se da la ley, ya lo haga esto la totalidad dicha o su parte prevalente por sí inmediatamente, ya lo haya encomendado hacer a alguno o algunos, que nunca son ni serán absolutamente hablando el legislador, sino sólo para algo y para algún tiempo y según la autoridad del primero y propio legislador. Y digo consiguientemente que por la misma autoridad primera, no otra, deben las leyes, y cualquiera otra cosa instituida por elección, recibir la aprobación necesaria, sea lo que sea de las ceremonias o solemnidades, que no se requieren para el ser de los elegidos, sino para su bien ser, el cual ceremonial, si se omitiere, no sería por ello menos válida la elección. Mas, de la misma autoridad deben las leyes y las otras cosas establecidas por elección recibir añadidos, supresiones o total mutación, interpretación o suspensión, según las exigencias de los tiempos, lugares y demás circunstancias, en las cuales fuere oportuno algo de eso por la común utilidad. Con la misma autoridad deben promulgarse o proclamarse las leyes después de su institución, a fin de que no pueda ningún ciudadano ni ningún forastero delincuente excusarse por su ignorancia.

§ 4. Llamo ciudadano, según Aristóteles, 3º de la *Política*, caps. 1º, 3º y 7º², a aquel que en la comunidad civil participa del gobierno consulti-

¹ *Polit.*, 1.2, c. 11; 1281 a 39.

² *Polit.*, 1. 3, c. 1; 1275 a 1; 1. 3, c. 3; 1277 b 33; 1. 3, c. 12; 1282 b 2; 1. 3, c. 11; 1281 a 40.

vo o judicial según su grado. Por esta delimitación quedan fuera de la condición de ciudadano los niños, los esclavos, los forasteros y las mujeres, aunque por razones diversas. Los niños de los ciudadanos son ciudadanos en potencia cercana por sólo el defecto de la edad. La parte prevalente de los ciudadanos conviene fijarla con arreglo a las honestas costumbres de las comunidades civiles, o determinarla según la opinión de Aristóteles, en el 6º de *Política*, cap. 2³.

§ 5. Definido así el ciudadano y la multitud prevalente de los ciudadanos, vengamos a nuestra intención propuesta, a saber, demostrar que la autoridad humana de dar la ley pertenece sólo a la totalidad de los ciudadanos o a la parte prevalente de ellos, lo que intentaremos, primeramente, del siguiente modo. La autoridad absolutamente primera de dar o instituir leyes humanas es sólo de aquél del que únicamente pueden provenir las leyes óptimas. Esa es la totalidad de los ciudadanos o su parte prevalente, que representa a la totalidad; porque no es fácil o no es posible venir todas las personas a un parecer, por ser la naturaleza de algunos tarda de nacimiento, o desentonar por malicia o ignorancia personal de la común opinión, por cuya irracional contestación u oposición no debe impedirse u omitirse lo útil a todos. Pertenece, pues, únicamente a la totalidad de los ciudadanos o a su parte prevalente la autoridad de dar o instituir las leyes.

La primera proposición de esta demostración es muy próxima a las por sí evidentes, aunque su fuerza y última certeza puede tomarse del cap. V de esta Parte. Pruebo la segunda proposición, a saber, que la ley óptima sólo sale de la auscultación y del precepto de toda la multitud, suponiendo, con Aristóteles, 3º de la *Política*, cap. 7º, que la mejor ley es la que se da para la utilidad común de los ciudadanos. De donde dijo: *Lo recto es*, se entiende, de las leyes, *de seguro lo tocante a lo útil para la ciudad y el común de los ciudadanos*⁴. Y que esto se hace del mejor modo sólo por la totalidad de los ciudadanos o por su parte prevalente, que se toma como una misma cosa con aquélla, lo muestro así; de aquello se juzga mejor la verdad y se advierte más diligentemente la común utilidad, a lo que toda la universalidad de los ciudadanos aplica su entendimiento y su afecto. Y más en condición está de advertir un defecto en la ley que se va a proponer y establecer la gran muchedumbre que cualquiera de sus partes, como *toda totalidad*, al menos la corpórea, en la mole y en la fuerza es mayor que cualquiera de sus *partes* tomada por separado. Además, toda la multitud atiende más a la utilidad común de la ley, porque nadie se daña a sí mismo

³ *Polit.*, l. 6, c. 3-4; 1318 a 4.

⁴ *Polit.*, l. 3, c. 13; 1283 b 40.

a sabiendas. Cada uno podrá ver allí si la ley propuesta se inclina más al bien de alguno o de algunos que al de otros o de la comunidad, y contra eso protestar; lo que no se haría si la ley fuera dada por uno solo o por pocos más atentos a su bien particular que al común. Reforzaré esta opinión lo que de la necesidad de la ley señalamos en el cap. XI de esta Parte.

§ 6. Todavía, volviendo a la conclusión principal; de aquél ha de ser solamente la autoridad de dar leyes, por el que a aquéllas mucho más perfecta y absolutamente se les da cumplimiento. Esto es sólo la universalidad de los ciudadanos, de ella es por tanto la autoridad de dar leyes. La primera proposición es muy próxima a las por sí evidentes, pues ociosa sería la ley si no se cumpliese. De donde Aristóteles, 4° de la *Política*, cap. 7°: *No es buena disposición de las leyes el poner bien las leyes y luego no obedecerlas. El mismo en el 6° de la misma, cap. 6°: Ningún provecho, dice Aristóteles, en el pronunciar sentencias sobre lo justo, pero no llevarlas a término*⁵. La segunda proposición la pruebo; porque aquella ley mejor se cumple por cualquiera de los ciudadanos que parece ser la que cada cual se ha impuesto a sí mismo; tal es la ley dada después de la auscultación y precepto de la multitud entera de los ciudadanos. La primera proposición de este silogismo se muestra casi evidente por sí misma, pues siendo la ciudad la comunidad de los hombres libres, como se escribe en el 3° de la *Política*, cap. 4°, todo ciudadano debe ser libre y no tolerar el *despotismo* de otro, es decir, un dominio servil⁶. Y ello no ocurrirá si la ley la diera alguno o algunos solos con su propia autoridad sobre la universalidad de los ciudadanos; dando así la ley serían déspotas de los otros. Y por eso los restantes ciudadanos, es decir, la mayor parte, llevarían pesadamente o de ningún modo la tal ley, por muy buena que fuera, y protestarían de ella víctimas del desprecio y, no convocados a su proclamación, de ningún modo la guardarían. Pero la dada con la audición y el consenso de toda la multitud, aun siendo menos útil, fácilmente cualquier ciudadano la guardaría y la toleraría, porque es como si cada cual se la hubiera dado a sí mismo y por ello no le queda gana de protestar contra ella, sino más bien la sobrelleva con buen ánimo. Todavía pruebo la segunda proposición del primer silogismo por otro argumento así: a aquél pertenece el poder de hacer observar la ley que tiene exclusivamente el poder coactivo contra los trasgresores; pero eso lo es la totalidad o la parte prevalente de ella, luego de ella sola es la autoridad de dar las leyes.

⁵ *Polit.*, 1. 6, c. 8; 1322 a 5.

⁶ *Polit.*, 1. 3, c. 6; 1279 a 21.

§ 7. Todavía al asunto principal arguyo así: aquello práctico en cuya debida institución consiste la mayor parte de la suficiencia de vida de los ciudadanos en este mundo, y de cuyo mal establecimiento amenaza un mal común, ha de establecerse sólo por la universalidad de los ciudadanos; eso es la ley, luego a la universalidad de los ciudadanos pertenece exclusivamente su institución. La mayor de esta demostración es próxima a las por sí evidentes y se sitúa entre las verdades inmediatas que se pusieron en los caps. IV y V de esta Parte. Pues se juntaron los hombres en una mutua comunicación civil para su utilidad y para conseguir una suficiencia de vida y evitar lo contrario. Mas lo que toca a la conveniencia o inconveniencia de todos, por todos debe ser conocido y oído para que puedan alcanzar lo conveniente y rechazar lo opuesto. Tales son las leyes, como se asumía en la menor. Porque en ellas, bien establecidas, consiste gran parte de toda la común suficiencia de vida humana, y en las inicuas la servidumbre, la opresión y la miseria, intolerables para los ciudadanos, de lo que finalmente viene la ruina a la comunidad política.

§ 8. Todavía, como compendio y suma de las anteriores demostraciones: o la autoridad de dar leyes pertenece a la totalidad de los ciudadanos, como dijimos, o a uno solo o a pocos. No a uno solo, por lo que dijimos en el cap. XI de esta Parte y en la primera demostración que adujimos para ello; podría efectivamente por ignorancia o por malicia o por ambas cosas dar una ley mala, mirando más al propio provecho que al común, por lo que sería tiránica. Por la misma causa tampoco pertenece a unos pocos; podrían, en efecto, pecar al dar la ley lo mismo que antes, con la mirada puesta, no en el bien común, sino en el de pocos, como se echa de ver en las oligarquías. Pertenece, pues, a la totalidad de los ciudadanos o a la parte prevalente, para lo que vale otra y contraria razón. Pues dado que todos los ciudadanos deben medirse con la ley en la proporción debida, y nadie a sabiendas se daña a sí mismo ni quiere para sí lo injusto, por ello todos, o los más, quieren la ley conveniente para la utilidad de todos los ciudadanos.

§ 9. Por las mismas razones queda demostrado que la aprobación, interpretación, suspensión de las leyes y las demás cosas propuestas en el párrafo 3 de este capítulo pertenecen a sola la autoridad del único legislador. Y lo mismo habrá que admitir de todo aquello que se establece por elección. Porque el que tiene la autoridad primera de elegir, él mismo aprueba o reprueba, o bien aquél a quien él le otorgare el poder de elegir; si no la parte sería mayor que el todo, o al menos igual a él, si lo que fue establecido por el todo pudiese ella anularlo.

XIII. DE ALGUNAS OBJECIONES A LO DICHO EN EL CAPÍTULO PRECEDENTE Y SUS
REFUTACIONES. UNA MÁS AMPLIA EXPOSICIÓN DEL INTENTO

Empero en lo que hemos dicho podrá alguno dudar disputando que no pertenece a la totalidad de los ciudadanos la autoridad de la legislación o de la institución. Primero, porque lo que por regla general es malo y falto de juicio no debe establecer la ley. Pues éstos son los dos defectos que deben excluirse en el legislador, a saber, la malicia y la ignorancia, y para evitarlo en los juicios, pusimos la necesidad de la ley en el XI de esta Parte. Pero tales son el pueblo o la totalidad de los ciudadanos; porque los hombres, por la mayor parte, son malos y necios. Pues *de los necios es infinito el número*, como tenemos en el *Eclesiastés*, 1^o¹. Además, porque es muy difícil o imposible concertar los pareceres de muchos malos e insensatos, lo que no ocurre tratándose de pocos y virtuosos. Más útil, pues, que la ley se dé por pocos que por la totalidad de los ciudadanos o excesiva pluralidad de ellos. Además, en cada comunidad civil hay pocos sabios y doctos comparados con la otra multitud indocta. Siendo, pues, más útil que la ley se dé por sabios y doctos que por indoctos y rudos, parece que la autoridad de darlas pertenece a pocos y no a muchos o a todos. Todavía, en vano se hace por muchos lo que por pocos puede hacerse. Pudiendo, pues, como se dijo, ser dada la ley por sabios, que son pocos, en vano se ocupará de ella la universal multitud o la mayor parte de ella. Así pues, no pertenece a la universalidad o a su parte prevalente la autoridad de dar la ley.

§ 2. De lo supuesto por nosotros anteriormente como principio de casi todo lo que vamos a demostrar en este libro, a saber, que *todos los hombres apetecen la suficiencia de vida y rechazan lo contrario*, sacamos como conclusión la comunicación civil de los mismos en el cap. IV de esta Parte, porque por ella obtienen aquella suficiencia de vida y sin ella no; por lo cual también Aristóteles en el 1^o de la *Política*, cap. 1^o, dijo: *Por naturaleza hay en todos un impulso a esa comunidad*, se entiende civil². De la cual verdad necesariamente se sigue otra y consta en el 4^o de la *Política*, cap. 10^o, a saber, que *es preciso que la parte que quiere que permanezca la sociedad civil prevalezca sobre la parte que no lo quiere*³, porque nada desea la misma naturaleza en su mayor parte y de un modo inmediato que signifique al mismo tiempo su disolución, pues sería en vano un tal deseo. Más aún, los que no quieren la permanencia de la sociedad civil, se cuentan entre los esclavos, no entre los ciudadanos, como son algunos advenedizos,

¹ *Eccl.*, 1. 15.

² *Polit.*, l. 1, c. 2; 1253 a 29.

³ *Polit.*, l. 4, c. 12; 1296 b 14.

de donde Aristóteles en el 7º de la *Política*, cap. 12º: *Con los sometidos se juntan todos los que hay por la región con ganas de subvertir*. Y añade: *Y que sean tantos en número los que no quieren vivir civilmente que prevalezcan sobre aquellos otros todos*, a saber, los que quieren la sociedad, *es una cosa imposible*⁴. Y que sea una cosa imposible, es manifiesto, porque esto sería errar o fallar la naturaleza en la mayor parte de los casos. Si, pues, la multitud predominante de los hombres quiere que permanezca la sociedad civil, como parece que ha quedado bien establecido, quiere también aquello sin lo cual la sociedad no puede conservarse. Esto es la regla de lo justo y lo útil dada como precepto, llamado ley, porque *es imposible que la ciudad aristocrática*, es decir, gobernada por la virtud, *no esté bien ordenada por las leyes*⁵, como consta en el 4º de la *Política*, cap. 7º, y lo demostramos en el XI de esta Parte. Quiere, pues, la multitud prevalente de la ciudad la ley, o se daría una quiebra en la naturaleza y el arte en la mayoría de los casos, lo que no se admite como posible por la ciencia natural.

Asumo además con todas las verdades anteriores manifiestas esta noción común, a saber, que *el todo es mayor que la parte*, lo que es verdadero, tanto en la magnitud o mole como también en la virtud activa y en la acción. De donde con bastante evidencia se sigue por necesidad que la universalidad de los ciudadanos o multitud prevalente, que se han de tomar como equivalentes, puede discernir sobre lo que se ha de elegir o rechazar, mejor que una cualquiera de sus partes.

§ 3. Esto, pues, supuesto como verdades comprobadas, fácil es rebatir las objeciones con las que alguien se empeñara en persuadir que el dar las leyes no pertenece a la totalidad de los ciudadanos ni a su multitud prevalente, sino a algunos pocos. Cuando se decía lo primero, que al malo y al torpe no le pertenece el poder de legislar, eso se concede. Y cuando se añadía que eso son la universalidad de los ciudadanos, se niega. Porque la mayoría de los ciudadanos ni es mala ni privada de discernimiento en los más de los casos supuestos y por mucho tiempo, porque todos o los más son de buen juicio y razón y de justa apetencia de la vida en sociedad política y de lo necesario para su permanencia y duración, como son las leyes y otros estatutos y costumbres, como ya se ha dicho. Porque aunque no pueda cualquiera ni la mayor parte de los ciudadanos inventar las leyes, puede, sin embargo cualquiera juzgar de las inventadas y las propuestas a él por otro, discernir si algo hay que añadir, quitar o cambiar. Por eso, si lo que en la proposición mayor se dice, *privada de discernimiento*, se entiende

⁴ *Polit.*, 1. 7, c. 14; 1332 b 29.

⁵ *Polit.*, 1. 4, c. 8; 1294 a 1.

que lo que por sí no puede inventar la ley en muchas de sus partes o supuestos, no debe establecer la ley, hay que negarla, como manifestamente falsa, valiendo como testimonio una inducción sensata y el 3° de la *Política* de Aristóteles, cap. 6°; por inducción, porque muchos juzgan rectamente de la cualidad de una pintura, o de una casa o de una nave y de los demás artefactos, aunque ellos no sepan inventarlos. Viene también a ello el testimonio de Aristóteles, en lo citado más arriba, respondiendo a la objeción con estas palabras: *Y porque de algunas cosas no será el que las hizo ni el único ni el mejor juez*⁶, aduciendo esto de muchos géneros de artefactos y dejando entender lo mismo de los demás.

[...]

§ 8. Y por ello es conveniente y sobre manera útil que las reglas, leyes futuras y estatutos de lo justo, útil y nocivo, lo que toca a las cargas comunes y cosas semejantes, el buscarlas o descubrirlas y examinarlas, se encomiende a los prudentes y expertos por la totalidad de los ciudadanos, de modo que, o bien separadamente, por cada una de las primeras partes de la ciudad, enumeradas en el V de esta Parte, párrafo 1, según la proporción de cada una, se elijan algunos, o bien por todos los ciudadanos congregados juntamente se seleccionen los varones expertos y prudentes predichos. Y éste será el modo conveniente y útil de congregarse para la invención de la ley sin hacer agravio a la restante multitud, a saber, de los menos doctos, que aprovecharía poco en el buscar esas reglas y sería perturbada en sus otros trabajos necesarios para sí y para los demás, lo que resultaría oneroso tanto para los particulares como para el común. Pero encontradas y diligentemente examinadas tales reglas, futuras leyes, deben ser propuestas en la asamblea de todos los ciudadanos reunidos para su aprobación o reprobación, de forma que si alguno de ellos le pareciere que hay algo que añadir, quitar, mudar o totalmente reprobear, pueda decirlo, porque por aquí podrá la ley más útilmente ordenarse. Pues, como hemos dicho, pueden los ciudadanos menos instruidos percibir alguna vez algo que corregir en la ley propuesta, bien que ellos fueran incapaces de descubrirla, porque así, dadas con la auscultación y consenso de la universa multitud, mejor se observarán y nadie podrá protestar contra ellas.

Hechas así públicas las reglas, leyes futuras, en la asamblea general, y oídos los ciudadanos que razonablemente quisieran decir algo sobre ellas, se habrán de elegir de nuevo varones de la condición y según el modo que dijimos, o confirmar los predichos, los cuales representando las veces y la autoridad de la totalidad de los ciudadanos, aprobarán o desaprobarán en

⁶ *Polit.*, 1. 3, c. 11; 1282 a 17.

todo o en parte las sobredichas reglas elaboradas y propuestas; o bien hará esto mismo; si lo quiere, la asamblea general de los ciudadanos por junto o su parte prevalente. Y ya después de esta aprobación las dichas reglas son leyes y merecen llamarse así, no antes; y ellas, después de su aprobación y proclamación, son las solas que, entre los preceptos humanos, obligan a los trasgresores bajo culpa y pena civil.

Por lo dicho creemos dejar demostrado suficientemente que la autoridad para dar o instituir leyes y para dar precepto coactivo de guardar las pertenece únicamente a la totalidad de los ciudadanos o a su parte prevalente, como a su causa eficiente, o a aquél o aquéllos a los que la misma totalidad se lo hubiere concedido.

XIV. DE LAS CUALIDADES O CONDICIONES DEL GOBERNANTE PERFECTO,
PARA SABER CÓMO TIENE QUE SER EL QUE HA DE SER ELEVADO AL PRINCIPADO.

DE DONDE TAMBIÉN SE DEDUCE LA MATERIA CONVENIENTE, ES DECIR,
EL SUJETO DE LAS LEYES HUMANAS

[...]

§ 4. Así, pues, por la prudencia se rigen las decisiones en los asuntos prácticos, por donde en el 6° de la *Ética*, cap. 4°, dijo Aristóteles que la prudencia es un hábito verdadero operante según razón en lo tocante a lo bueno y malo del hombre¹, se entiende en cuanto hombre. Y la causa de esto es que, siendo los asuntos prácticos aquéllos sobre los cuales frecuentemente tratan las leyes humanas, según las cuales debe el gobernante ordenar los actos civiles humanos, parece que no todos sus modos o circunstancias en que vienen envueltos pueden siempre de una vez determinarse por la ley, por su variedad y diversidad, atendiendo a las regiones y a los tiempos, como lo enseña la experiencia patente y testigo de ello es Aristóteles en el 1° de la *Ética*, cap. 1°, cuando dijo: *Lo bueno y lo justo, de lo que la ciencia civil trata, son de condición tan diversa y voluble, que parece que sólo existen por ley y no por naturaleza*², es decir, porque los hombres quieren establecerlo así, y no porque la misma naturaleza de esos asuntos prácticos lo haya así determinado, que esto sea justo, aquello injusto. Y lo mismo, más ampliamente declarado, lo tiene en el 3° de la *Política*, cap. 9°, cuando dice: *Y porque unas cosas pueden ser comprendidas en las leyes y otras es imposible, estas últimas hacen dudar y preguntarse si es preferible que gobierne una ley óptima o un varón óptimo. Las cosas sobre las que*

¹ *Eth. Nico.* 1. 6, c. 5; 1140 b 5.

² *Eth. Nico.* 1. 1, c. 3; 1094 b 14.

deliberan, se entiende, los hombres, *es imposible que estén determinadas por la ley*, suple, todas³.

§ 5. Por esto convino que el juicio de algunas cosas que acaecen en los actos civiles de los hombres se dejen al arbitrio de los gobernantes, cosas que en sí mismas o en alguno de sus modos o circunstancias no están determinadas por la ley. Pues en lo que está determinado por la ley el gobernante debe atenerse a la determinación legal.

[...]

§ 7. Y es también conveniente al futuro gobernante que adquiera la virtud que se llama *epikeya*, por la cual se guía el juez, mayormente según el afecto, en aquello, en lo que la ley es deficiente. De donde Aristóteles en el 4º de la *Ética*, tratado de la justicia, dice: *Y ésta es la naturaleza de la que se llama epieikes, es decir, un modo directivo de la ley, donde ésta falla por razón de un caso particular*⁴.

[...]

§ 8. Fuera de estos hábitos y disposiciones dichas es necesario al gobernante algún instrumento externo, a saber, un cierto número de hombres armados con los cuales pueda poner en ejecución sus sentencias civiles contra los rebeldes y desobedientes por medio de un poder coactivo. [...]

§ 9. En todo caso, reduciendo a capítulo lo que toca a las disposiciones y demás cosas necesarias a los gobernantes, digamos que antes de la elección se necesitan prudencia y virtud moral en el que ha de ser elegido, o los que han de ser elegidos, si son muchos, como ocurre en la aristocracia. El poder armado es necesario al que tiene el supremo mando en la ciudad o en el reino, como instrumento o aparato externo para que se cumplan sus sentencias según las leyes; pero no conviene que tenga ese poder antes de su elección, sino que se le otorgue juntamente con el gobierno. El amor o la benevolencia singular para con la comunidad política y para con los ciudadanos contribuye a la bondad y solicitud de sus acciones civiles aunque no se exige de él con la misma necesidad que las antedichas cualidades.

[...]

XV. DE LA CAUSA EFICIENTE DE LA MEJOR INSTITUCIÓN DEL GOBIERNO. DE DONDE TAMBIÉN SE DEDUCE LA CAUSA EFICIENTE DE LAS OTRAS PARTES DE LA CIUDAD

[...]

Perteneciendo, pues, a la totalidad de los ciudadanos engendrar la forma según la cual los actos civiles todos deben regirse, es decir, la ley, a la misma totalidad pertenece determinar la materia de esta forma, o sea, el

³ *Polit.*, I, 3, c. 16; 1287 b 19.

⁴ *Eth. Nic.*, I, 5, c. 10; 1137 b 27.

sujeto al cual toca disponer los actos civiles de los hombres según aquella forma, es decir, la parte gobernante. y como ésta es la suprema y óptima de las formas de la comunidad civil, es preciso que se le determine el mejor sujeto en cuanto a disposiciones; lo cual también lo concluimos con razones probables en el capítulo precedente. De donde parece deducirse convenientemente que para un régimen político se ha de preferir el gobernante elegido, y absolutamente sin derecho de sucesión hereditaria, a los no elegidos, o sea, los que se instituyen con sucesión hereditaria.

§ 4. Una vez expuesta la causa eficiente de esta parte de la ciudad, nos cumple decir, según lo propuesto muchas veces por nosotros, la causa eficiente, instituyente y determinante de los otros oficios o partes de la ciudad. A la primera la llamamos legislador, a la segunda, como instrumental o ejecutiva, llamamos gobernante por la autoridad a ella concedida por el legislador, según la forma dada por el mismo, es decir, la ley, con arreglo a la cual ha de obrar y disponer siempre en lo posible en los actos civiles, como hemos mostrado en el capítulo precedente. Porque aunque el legislador, como primera y apropiada causa, es quien ha de determinar por quiénes y cuáles oficios hayan de ejercitarse en la ciudad, la ejecución de los tales, como de las demás cosas legales, es la parte gobernante la que las prescribe y si es preciso las prohíbe. Es, en efecto, más hacendera la ejecución de lo legal por ella que por la multitud de los ciudadanos, porque para eso basta uno o pocos gobernantes, en lo que en vano se ocuparía la universalidad de la comunidad, que se vería distraída de otras labores necesarias. Y que haciéndolo aquéllos lo hace toda la comunidad, porque lo hacen los gobernantes de acuerdo con la determinación de la comunidad, a saber, la determinación legal, y con pocos o uno solo es más fácil la ejecución de lo legal.

[...]

§ 13. Y por ello nunca debe cesar en la ciudad la acción del gobernante, como ni la acción del corazón en animal. Porque aunque las acciones de las otras partes de la ciudad puedan cesar por algún tiempo sin daño para ningún particular, grupo o comunidad, como la acción de la parte militar en tiempo de paz, y así semejantemente de otras partes, pero la acción primera de esta parte gobernante y de su virtud nunca puede cesar sin daño. Porque en cualquier hora o momento ha de estar en pie el precepto y la común custodia de lo lícito y de lo ilícito según la ley, y cuando quiera que se produzca lo ilícito o lo injusto, conviene que la parte gobernante lo regule perfectamente o que ponga en acción aquellas cosas que son previas a su regulación.

§ 14. De lo dicho puede quedar suficientemente claro el orden de las partes de la ciudad entre sí, y que por el gobernante y por relación a él, como a lo primero de todo, se ordenan todas en el estado de la vida

presente. Pues en la comunidad civil es la primera aquella parte que ha de instituir las demás, determinarlas y conservarlas en su estado, en este mundo, es decir, para el fin civil; y ésta es la parte gobernante según la ley humana, como lo hemos concluido con la razón probable y demostrativa. Es, pues, la primera de todas las otras y a ella se ordenan las demás.

Quede así determinado sobre la causa eficiente de la elección de la parte gobernante, así como de la institución de las otras partes de la ciudad y su orden y acciones recíprocas.

[...]

XVIII. DE LA CORRECCIÓN DEL PRÍNCIPE Y POR QUÉ CAUSAS, CÓMO Y POR QUIÉNES DEBE SER CASTIGADO SI ES TRASGRESOR DE LA LEY

Hemos dicho en lo que precede que las correcciones o totales cambios de los gobernantes pertenecen al legislador igual que la institución de ellos. En lo cual alguien podría dudar si es conveniente sancionarlos por un juicio y un poder coactivo. Y si se ha de hacer, si por cualesquiera abusos o por algunos sí y por otros no. También a quién o a quiénes pertenezca hacer tales juicios y cumplir las sentencias de los juicios con poder coactivo, habiendo dicho anteriormente que a solos los gobernantes corresponde el dictar las sentencias civiles y castigar a los trasgresores de las leyes con poder coactivo.

§ 2. Digamos que el príncipe, por su acción según la ley y la autoridad a él otorgada, es la regla y la medida de cualquier acto civil, como el corazón en el animal, según se ha mostrado en el XV de esta Parte. Si el príncipe no recibiera otra forma que la de la ley, la autoridad y el deseo de obrar en consecuencia, nunca realizaría una acción sancionable o medida por otra cosa. Y así, tanto él como su acción, de tal manera sería la medida de cualquier acto civil de los demás, que nunca se mediría por otros. Como el corazón bien formado en el animal, el cual, al no recibir otra forma que le incline a una acción contraria a la acción que proviene de su forma y calor natural, siempre realiza la acción que conviene naturalmente y nunca la contraria, por lo cual, de tal manera regula y mide con su influencia y acción las otras partes del animal, que nunca es regulado por éstas ni recibe su influencia.

§ 3. Pero como el príncipe, por ser hombre, tiene su entendimiento y su apetito capaces de recibir otras formas, como, por ejemplo, una falsa apreciación o un deseo perverso, o ambas cosas, en fuerza de las cuales ocurre que él obre de modo contrario a lo que prescribe la ley, por ello el príncipe respecto de esas acciones se hace mensurable por otro que tenga autoridad de medir y regular según la ley a él o a sus acciones trasgresoras

de la ley; de otro modo todo gobernante se tornaría despótico y la vida civil servil e insuficiente. Lo que es un inconveniente que hay que evitar, como se deduce de lo que hemos expuesto en el V y XI de esta Parte.

Pero el juicio, el precepto y la ejecución de cualquier corrección del príncipe, según su demérito o trasgresión, debe hacerse por el legislador o por alguno o algunos constituidos en autoridad para ello por el legislador, como demostramos en el XII y XV de esta Parte. Conviene también por algún tiempo suspender el oficio del príncipe que ha de ser sancionado, por respeto, sobre todo, a aquel o a aquellos que deberán juzgar su trasgresión, no sea que por la pluralidad del gobierno se produzca en la comunidad un cisma, algarada o lucha, y porque no es juzgado en cuanto príncipe, sino en cuanto súbdito trasgresor de la ley.

XIX. DE LAS CAUSAS EFICIENTES DE LA TRANQUILIDAD DE LA CIUDAD O DEL REINO, Y DE AQUELLA QUE DE UN MODO INSÓLITO PERTURBA EN PARTICULAR A LOS REINOS

§ 12. [...] Porque no le pertenece al obispo romano, ni a otro obispo, sacerdote o cualquiera otro ministro sagrado, en cuanto tales, el cargo de gobierno coactivo sobre cualquier persona particular de cualquier condición que sea, comunidad o grupo, como lo hemos demostrado en el cap. XV y en el XVII de esta Parte. Y esto, de cualquier sacerdocio en cualquier religión o secta lo pensó Aristóteles en el 4º de la *Política* cuando dijo: *Por lo cual no todos los que son elegidos o sacados por suerte se han de poner como príncipes, por ejemplo, en primer lugar, los sacerdotes. Porque esto es algo que cae fuera de los cargos de gobierno políticos, etc. Unos están entre los cometidos políticos, es decir, los cargos, etc. Y poco después añade: Y otros son de carácter doméstico*¹. [...]

§ 13. Y porque esta parte perniciosa, enemiga total de la paz humana y su felicidad, puede inficionar en grado máximo, por el mismo vicio de la raíz corrompida, a todos los demás reinos de los fieles cristianos del mundo, juzgo necesarísimo rechazarla, como en los comienzos quedó dicho. Primero, rasgando la envoltura de aquella opinión dicha, como raíz de los males ya acaecidos y por acaecer; después, frenando con acciones externas, si es preciso, a sus adalides o inventores, ignorantes o injustos y a sus obstinados defensores. Urge que le hagan frente todos los que saben y pueden. Y los que lo descuidan u omiten, teniendo alguna oportunidad para ello, son injustos según el testimonio de Tulio, en el 1º del *De los deberes*, cap. 5º, cuando dice: *Dos géneros hay de injusticia, el de aquellos que la cometen, y el de aquellos que, pudiendo, no apartan la injusticia de quienes la padecen*². [...]

¹ *Polit.*, 1. 4, c. 15; 1299 a 16.

² *Cic.*, *De officiis*. I, 7.

PARTE SEGUNDA

I. DE LOS TRES IMPEDIMENTOS O MODOS DE CONTRADECIR A LAS VERDADES
CONTENIDAS EN LA PRIMERA PARTE; DEL SENTIDO Y ALCANCE DE LOS TEMAS
QUE SE HAN DE TRATAR Y DEL MODO DE PROCEDER

[...]

§ 3. Procederé en la Segunda Parte de este libro, primero, aduciendo las autoridades del sagrado canon con algunas imaginarias y extrañas interpretaciones de algunos, por las cuales parecería que es posible persuadirse de que al obispo de los romanos se le debe de derecho la suprema de todas las jurisdicciones coactivas o del principado, sobre todo en el ámbito de la ley cristiana, tanto sobre el príncipe romano, como sobre todos los otros principados, comunidades, grupos y personas particulares, aun las seglares, y mucho más sobre los presbíteros u obispos, diáconos, y sus colegios y personas particulares de cualquier condición que fueren. Porque si con algún género de necesidad se deduce, en virtud de las palabras de la Escritura, que vale eso para uno de esos grupos, seculares o clérigos, con la misma necesidad se concluirá para todos los demás.

A éstas añadiré después algunas razones de orden político para apoyar el susodicho propósito, razones que, al pronto, se ofrecen como muy claras, supuestas ciertas verdades de la Sagrada Escritura, que, bien concertadas entre sí, han de ser supuestas por todos como verdaderas. Las aduciré, digo, a fin de refutarlas a toda luz, para que, en virtud de ellas, o de otras similares, nadie pueda ya llamarse a engaño, y para que de su refutación quede más claro lo débil de la opinión antedicha, a la que por mucho tiempo hasta ahora se le dio visos de verosimilitud.

§ 4. Después de esto aduciré, haciendo contraste con lo anterior, verdades sacadas del mismo canon de los libros sagrados, con las exposiciones de los santos intérpretes, en nada imaginarias, ni extrañas ni falsas, sino concordes entre sí y apropiadas, con las que se demuestra abiertamente que ni el obispo romano llamado Papa, ni ningún otro cualquier presbítero, obispo o ministro espiritual, individual o colectivamente, en cuanto tales, ni tomados en grupo como colegio, tiene ni debe tener jurisdicción alguna real o personal coactiva sobre cualquier presbítero, obispo o diácono o sobre su colegio; y que mucho menos el mismo o alguno de ellos colegialmente o individualmente tiene tal jurisdicción sobre cualquier príncipe o principado, comunidad, colegio o persona particular seglar, de cualquier condición que fuere, a no ser que en último término esa tal jurisdicción le hubiere sido concedida por el legislador humano en determinada provincia a un presbíte-

ro, o a algún obispo o a su colegio. Para demostrar lo cual y confirmarlo, podrán y deberán aducirse, cuando fuere oportuno para alguno, las pruebas políticas, cuyos fundamentos propios se contiene en los caps. XII, XV y XVII de la Primera Parte. Pues repetirlas aquí no es nuestro propósito por razón de brevedad. De acuerdo con ello mostraré cuál y cuánta es la potestad sacerdotal y de las llaves dada por Cristo al obispo romano y a cualquier otro sacerdote. Pues de la clarificación de estas cosas depende la solución de muchas dudas tocantes a la verdad y abrirán camino al fin que nos proponemos alcanzar.

§ 5. Después, será altamente útil pronunciarnos sobre algunas cuestiones implicadas en lo dicho; en conexión con las cuales diremos algo oportuno para nuestro propósito referente a los privilegios ya concedidos por los príncipes romanos a los obispos romanos. Porque de aquí tomaron su origen ocasionalmente las usurpaciones u ocupaciones, o detenciones de jurisdicciones coactivas, que atribuyen ahora a su autoridad propia los obispos de los romanos, y que por la costumbre, mejor, el abuso, cobraron luego su fuerza, máxime durante la sede imperial vacante. Porque en estos privilegios desde el principio, no de otras procedencias, se han apoyado los obispos de los romanos para la adquisición y conservación de esas jurisdicciones coactivas. Andando el tiempo, o bien porque fueron privados de ellas por sus deméritos, o para que no aparezca lo débil de esa circunstancia histórica dicha y lo patente de la indebida jurisdicción ya por ellos asumida, para solapar su ingratitud, o bien, acaso, como dejamos expuesto como verosímil en el último capítulo de la Primera Parte, dado que pretenden la ocupación de la jurisdicción coactiva de todos los reinos para la que ningún apoyo tienen en los mencionados privilegios, el hecho es que ya no echan mano de ellos, sino que se sirven de otro universal para someter a sí a los príncipes y a todos los que viven en una comunidad civil, a saber, el título supradicho de la plenitud de potestad. En razón de la cual, otorgada por Cristo al bienaventurado Pedro, como a su vicario, cualquier obispo romano proclama poseer, en cuanto sucesor de Pedro, la jurisdicción coactiva suprema sobre todos los hombres y provincias.

Lo último de esta Parte será mostrar que las autoridades del sagrado canon escriturístico, que parecen contradecir a la opinión que juzgamos acorde con la verdad y con la Escritura, en nada apoyan el error predicho, sino más bien lo rebaten, como aparecerá evidente a la luz de las exposiciones de algunos santos y no menos de las de otros aprobados doctores de la fe cristiana, por las cuales aparecerá también que las exposiciones, mejor diremos fantasías, de algunos enemigos empeñados en trastocar la Escritura para apuntalar el sentido de la falsa apreciación antedicha, son violentas,

extrañas a la Escritura y distorsionadas, y en disonancia con la idea de los santos y de los doctores avisados de la fe cristiana.

Lo último de todo será rebatir las razones que llamo de orden político, que ofrecen la apariencia de respaldar la falsa apreciación tantas veces mencionada.

II. DE LA DISTINCIÓN DE LOS TÉRMINOS O EXPRESIONES COMPRENDIDOS EN LAS CUESTIONES QUE VAMOS A TRATAR

Pero antes que entremos en la discusión de los temas propuestos, a fin de evitar que de la multiplicidad de los términos que usaremos en las cuestiones principales, surja alguna ambigüedad o confusión de las opiniones que vamos a exponer, distingamos el significado de los mismos. Porque como se dice en el libro primero de los *Elencos: Los ignorantes de los nombres caerán en paralogismos, como disertantes y como oyentes*¹. Los nombres o expresiones, cuya multiplicidad queremos distinguir, son éstos: *iglesia, juez, espiritual, temporal*. Porque de la inquisición propuesta queremos saber si pertenece al obispo romano o a cualquier otro obispo o presbítero o diácono, o al colegio de ellos, los que suelen decirse hombres de iglesia, eclesiásticos, ser juez coactivo de asuntos temporales o espirituales o de ambos, o si ni para los unos ni para los otros son ellos tales jueces.

§ 2. Abordando, pues, el tema, digamos que este nombre, *iglesia*, es un vocablo tomado del uso de hablar los griegos y significa para ellos, de lo que ha llegado hasta nosotros, la congregación del pueblo comprendida bajo un régimen, como tomó la palabra *iglesia* Aristóteles en el 2º de la *Política*, cap. 7º, cuando dice: *Y de la iglesia participan todos*².

Entre los latinos este vocablo, según la acepción vulgar y corriente en una de sus significaciones, importa el templo o la casa en la que comúnmente se da culto a Dios por los fieles y frecuentemente es adorado. Pues sobre la iglesia se expresa así el apóstol, en la primera a los Corintios, cap. 11º: *¿Es que no tenéis casas para comer y beber o hacéis menos la iglesia de Dios?*³ Donde la glosa, según Agustín: *“Hacéis menos la iglesia”, es decir, la casa de oración*, y poco más abajo añade: *Esto trajo consigo el uso cotidiano, que “el entrar o refugiarse en la iglesia” no se dice sino de aquél que entra o se refugia en el mismo lugar dentro de sus muros*⁴.

En otra significación se comprenden bajo este nombre *iglesia* todos los presbíteros u obispos, diáconos y demás ministros del templo o de la

¹ *Soph Elench.*, I, 165 a 15.

² *Polit.*, I, 2, c. 10; 1272 a 10.

³ *I Cor.*, 11, 22.

⁴ P. Lomb., *Collectanea in epistolas sancti Pauli*. PL 191,1639 C.

iglesia tomada según la primera significación. Y según este sentido se dicen corrientemente *iglesia* las personas eclesiásticas, varones eclesiásticos, sólo los clérigos o ministros antedichos.

Todavía en otra significación, máxime entre los modernos, bajo el nombre *iglesia* se comprenden aquellos ministros, presbíteros u obispos y diáconos, que sirven y presiden en la metropolitana o principal de todas las iglesias, como ya de antiguo adquirió esta denominación la iglesia de la urbe romana, cuyos ministros y presidentes son el Papa romano y sus cardenales, que ya por cierto uso llegó a llamárseles *iglesia* y que la *iglesia* hizo o aceptó algo cuando aquéllos hicieron o recibieron algo, o de otro cualquier modo dispusieron algo.

§ 3. También, según otra significación este nombre *iglesia* en el más verdadero y propísimo sentido, según la primera imposición de este nombre o la intención de los primeros en imponerlo, aunque no tan corriente o de acuerdo con el uso moderno, se dice de la universalidad de los fieles que creen e invocan el nombre de Cristo, y de todas las partes de esta universalidad en cualquiera comunidad suya aun la doméstica. Y ésta fue la primera imposición de este vocablo y la acepción usual de él entre los apóstoles y en la iglesia primitiva. De donde el apóstol en la primera a los Corintios, cap. 1º: *A la iglesia que está en Corinto, a los santificados en Cristo Jesús, llamados santos, con todos los que invocan el nombre de Nuestro Señor Jesucristo*⁵. Donde la glosa, según Ambrosio: *A los santificados en el bautismo, y esto en Cristo Jesús*⁶. Y en este sentido hablaba el apóstol en los Hechos, cap. 20º a los presbíteros de Efeso cuando dijo: *Atended a vosotros y a la grey toda, en la que os puso a vosotros como obispos el Espíritu Santo, para que rijáis la iglesia de Dios que adquirió él con su sangre*⁷. Y por ello son y deben decirse, con verdaderísima y propísima significación, varones eclesiásticos todos los fieles cristianos, tanto los sacerdotes como los no sacerdotes, por razón de que a todos los adquirió y redimió Cristo con su sangre.

[...]

§ 4. A seguido conviene distinguir estos nombres, *temporal* y luego *espiritual*. De los cuales por sernos más familiar, comencemos diciendo que esta expresión *temporal* en una de sus significaciones más corrientes, se dice de todo lo corporal, natural o artificial, distinto del hombre, que, existiendo en algún modo bajo su potestad, se ordena a su uso, a subvenir a

⁵ *I Cor.*, 1, 2.

⁶ *Collect.*, PL 191, 1535 A.

⁷ *Hechos*, 20, 28.

su necesidad y a procurarle placer en el estado y para el estado de la vida en este mundo. Del cual modo más generalmente acostumbró a decirse temporal todo lo que comienza y cesa en el tiempo. Esto es lo que es y se dice propiamente ser en el tiempo, como en el 4º de la Física⁸.

De otro modo se dice *temporal* de todo hábito, acción o pasión humana en el mismo sujeto o con vistas a otro, llevado a cabo por el hombre mirando el fin de este mundo o de la vida presente.

Todavía, menos universalmente se dice este nombre *temporal* de las acciones y pasiones voluntarias y transitivas ordenadas al bien o daño de otro distinto del que las realizó, y de éstas máxime tratan los legisladores humanos.

§ 5. Ahora quiero distinguir los significados o sentidos de esta expresión, espiritual, que en una de sus acepciones se dice de todas las substancias incorpóreas y sus acciones.

En otra significación se dice de toda acción o pasión humana, de su potencia de conocer y apetecer inmanente a ella. Según esta significación se dicen también espirituales las acciones de las cosas corporales en los sentidos de los animales y sin materia, como los simulacros, imaginaciones, fantasmas y especies de las cosas, que son para el alma como la razón de conocer, y en ese género se ponen por algunos las acciones de las cosas sensibles, aun en substancia no animada, como las generaciones de la luz y de otras cosas parecidas.

De nuevo, y más para nuestro propósito, se dice este nombre *espiritual* de la ley divina, de la doctrina y de la enseñanza de los preceptos y de los consejos derivados de ella y para ella. En la cual significación entran también todos los sacramentos eclesiásticos y sus efectos, toda la gracia divina, las virtudes teológicas todas y los dones del Espíritu Santo ordenados para nuestra vida eterna. Pues así con propiedad se sirve de este nombre el apóstol a los Romanos, 15º y 1ª a los Corintios, 9º, cuando dice: *Si sembramos en vosotros cosas espirituales, ¿será mucho que seguemos de vosotros cosas carnales?*⁹. Donde la glosa, según Ambrosio: *Espirituales, es decir, aquello que vivifica vuestro espíritu, o lo que se os dio por el Espíritu Santo, a saber, el Verbo de Dios y el misterio del reino de los cielos*¹⁰.

Todavía, según otra significación, suele tomarse este nombre por cualquiera acción o pasión humana voluntaria, tanto en sí misma, como la llevada a cabo con vistas a otro, poniendo el deseo en el mérito para la vida

⁸ *Phys.*, 1. 4, c. 12; 221 a 3 s.

⁹ *Rom.*, 15, 27. *1 Cor.*, 9, 11.

¹⁰ *Collect.*, PL 191, 1609 B.

bienaventurada del siglo venidero; tales son los actos de contemplación de Dios, el amor del mismo y de los prójimos, las abstinencias, las obras de misericordia, de mansedumbre, oraciones, ofrendas inspiradas en la piedad o dirigidas al culto divino, los actos de hospitalidad, peregrinaciones, penitencias corporales, fuga y desprecio de deleites carnales, y generalmente todas las cosas semejantes a éstas ordenadas al fin antedicho.

Aún se dice este nombre, si no con tanta propiedad como el segundo y tercer modo, del templo o iglesia dicha según la segunda significación, de todos los vasos y utensilios que en él se ordenan al culto divino.

Pero de todas estas significaciones una última, no correcta e impropia en alto grado, es aquélla a la que extienden algunos este nombre para significar las acciones voluntarias transitivas y sus omisiones de parte de los presbíteros u obispos, diáconos y otros ministros del templo, encaminadas al provecho o daño de otro distinto del que realiza la acción, con vistas a la vida en este mundo.

Extienden, también, y aún más impropriamente, el mismo nombre a las posesiones y bienes temporales, muebles e inmuebles de los mismos y a ciertos censos y bienes temporales que llaman diezmos, para, con el pretexto de este vocablo, eximirse de las normas legales y las emanadas de los gobiernos civiles.

[...]

§ 7. Lo mismo habrá que pensar sin duda de ciertas acciones de los presbíteros, obispos y diáconos. Porque no son ni se deben decir espirituales todos los actos de ellos; muchos, en efecto, son actos civiles, contenciosos y carnales o temporales. Porque pueden dar préstamos, hacer depósitos, comprar, vender, pegar, matar, robar, cometer adulterio, arrebatar, traicionar, engañar, dar falso testimonio, difamar, caer en herejía y perpetrar otros crímenes, delitos y pendencias, igual que pueden ser cometidos por otros no sacerdotes. Por lo que habrá que preguntarles a ellos, a propósito, si sus acciones posibles que hemos reseñado son o pueden ser dichas espirituales por alguien de sana mente. Bien claro es que no lo son, sino carnales o temporales.

[...]

§ 8. Queda distinguir estos nombres *juez* y *juicio*, que expresa la acción de aquél. Porque son de aquéllos de múltiple apelación y en los que, al querer definir las cuestiones, ofrecen ambigüedad y obstáculo por su misma multiplicidad. Se dice *juez*, en una significación, de todo el que discierne y conoce, máxime con referencia a un hábito especulativo u operativo; y este nombre, *juicio*, del conocimiento o discernimiento de los tales. Del cual modo, el geómetra es juez y juzga de las figuras y de sus

peculiaridades, y el médico juzga de los sanos y de los enfermos, y el prudente de lo que hay que hacer u omitir, y el constructor de casas de cómo se han de construir. Así también el que sabe o es experto se dice juez y juzga de las cosas cognoscibles o realizables que le atañen. Y según esta significación tomó Aristóteles estos nombres en el 1º de la *Ética*, cap. 1º, cuando dijo: *Cada uno juzga bien en lo que conoce y de ello es buen juez*¹¹.

También se dice este nombre *juez*, con otra significación, del que tiene ciencia del derecho político o civil que, con apelación corriente suele decirse *abogado*, aunque en las más de las provincias y mayormente en la de los itálicos se denomina *juez*.

De otro modo también se dice este nombre *juez* del príncipe o gobernante, y el nombre de *juicio* de la sentencia del gobernante, cuya autoridad, en efecto, está precisamente en juzgar de lo justo y lo útil según las leyes o las costumbres y en dictar y ejecutar con poder coactivo las sentencias dadas por él mismo.

[...]

IV. DE LAS ESCRITURAS CANÓNICAS, DE LOS PRECEPTOS Y CONSEJOS Y EJEMPLOS DE CRISTO Y DE SANTOS Y PROBADOS DOCTORES EXPOSITORES DE LA LEY EVANGÉLICA.

CON LOS CUALES SE DEMUESTRA CON EVIDENCIA QUE EL ROMANO Y CUALQUIER OTRO OBISPO O PRESBITERO O CLÉRIGO SIN EXCEPCIÓN, EN VIRTUD DE LAS PALABRAS DE LA ESCRITURA, NINGÚN PRINCIPADO COACTIVO O JURISDICCIÓN CONTENCIOSA PUEDEN REIVINDICAR PARA SÍ O ATRIBUIRSE, Y MUCHO MENOS EL SUPREMO DE TODOS SOBRE CUALQUIER CLÉRIGO O LAICO; Y QUE, A TENOR DEL CONSEJO Y EL EJEMPLO DE CRISTO, MÁXIME EN LAS COMUNIDADES DE LOS FIELES. SI SE LES OFRECE O SE LES CONFIERE POR EL QUE TIENE AUTORIDAD PARA ELLO, DEBEN RECUSARLO, Y ADEMÁS, QUE TODOS LOS OBISPOS Y LOS QUE INDISCRIMINADAMENTE LLAMAMOS CLÉRIGOS, HAN DE SOMETERSE AL JUICIO COACTIVO O GOBIERNO DEL QUE MANDA CON AUTORIDAD DEL LEGISLADOR HUMANO, SOBRE TODO SI ES FIEL

En contra de ello queremos seguidamente aducir las verdades de la Escritura sagrada, mandando o aconsejando expresamente, ya sea en su sentido literal, ya en el místico, con las interpretaciones de los santos y con la exposición de la fe cristiana de otros aprobados doctores, que el obispo romano llamado Papa y cualquiera otro obispo, o presbítero o diácono, no tengan ni deban tener el gobierno o juicio o jurisdicción coactiva sobre cualquier sacerdote o no sacerdote, gobernante, comunidad, colegio o per-

¹¹ *Eth. Nic.*, l. 1, c. 1; 1094 b 28.

sona particular de cualquier condición, entendiendo por *juicio coactivo* aquel que en el cap. 11 de esta Parte dijimos entenderse según la tercera significación de juez y juicio.

§ 2. Mas para declarar con más evidencia este sentido no debe ocultársenos que en esta investigación no nos preguntamos cuál sea el poder o autoridad que Cristo, verdadero Dios y verdadero hombre, tuvo o tenga en este presente siglo, ni qué y cuánto pudo conferir de esto al bienaventurado Pedro y a los demás apóstoles y a sus sucesores, obispos o presbíteros, porque de esto no dudan los fieles cristianos en los asuntos aquí tratados. Sino queremos y debemos inquirir qué clase de poder y autoridad quiso Cristo conferirles para ejercerla en este siglo y de hecho les confirió, y de cuál los excluyó y se lo prohibió con el consejo o con el ejemplo. Porque estamos obligados a creer que ellos tienen de Cristo tanta potestad y autoridad cuanta podemos deducir de las palabras de la Escritura que les fue entregada, no otra. Porque es cierto para todos los fieles de Cristo, que Cristo, verdadero Dios y verdadero hombre, pudo conferir, no ya sólo a los apóstoles, sino a cualesquiera otros, autoridad o jurisdicción coactiva en este siglo sobre todos los príncipes o principados y todas las demás personas particulares del mundo, y quizá aún una mayor, como de crear entes, destruir o reconstruir el cielo y la tierra y lo que en ella hay, e incluso de mandar sobre los ángeles: lo cual, sin embargo, Cristo no se lo dio ni pensó en dárselo. De donde Agustín, en el Sermón 10º: *De las palabras del Señor, según Mateo*, dice así: *Aprended de mí, no a fabricar el mundo, no a crear todo lo visible y lo invisible, no a hacer milagros en el mismo mundo y resucitar muertos, sino que soy manso y humilde de corazón*¹.

§ 3. Por tanto, apuntando al fin que me he propuesto, basta mostrar y mostraré, lo primero, que Cristo mismo no vino al mundo a dominar a los hombres, ni a juzgarlos con el juicio según la tercera significación, ni a gobernar temporalmente, sino más bien a someterse dentro del estado del presente siglo; más aún, de tal juicio o gobierno en ese mismo sentido se excluyó a sí mismo y a sus apóstoles y discípulos suyos; y consiguientemente a los sucesores de los mismos, obispos o presbíteros, los excluyó con su ejemplo y con su palabra de consejo y de precepto de todo principado o gobierno terreno semejante. Lo mismo mostraré que hicieron los apóstoles principales como verdaderos imitadores de Cristo y enseñaron a sus sucesores a hacer lo mismo; más aún, que tanto Cristo como los apóstoles mismos quisieron someterse y se sometieron constantemente a la jurisdicción coactiva de los gobernantes del mundo, real y personalmente, y a todos los

¹ PL 38, 441.

demás, a los que predicaron la ley de la verdad o se la legaron por escrito, enseñaron y mandaron hacer lo mismo bajo pena de condenación eterna. Después pondré un capítulo sobre la potestad o autoridad de las llaves otorgada por Cristo a los apóstoles y a sus sucesores en el oficio, obispos o presbíteros, a fin de que quede en claro qué, cuál y cuánta sea esa potestad, tanto en el obispo romano como en los otros. Porque hubo siempre ignorancia de ello, pero ahora es el origen de muchas discusiones y perniciosas reyertas entre los fieles de Cristo, como en el primer capítulo de esta Parte de algún modo apuntábamos.

§ 4. Prosiguiendo, pues, en nuestro intento queremos mostrar que Cristo, consecuente con su propósito e intención, de palabra y de obra, se quiso excluir y se excluyó a sí mismo y a sus apóstoles del oficio de gobernar o de la jurisdicción contenciosa, del régimen o juicio coactivo cualquiera en este mundo. Esto aparece sin lugar a duda, primero, en el pasaje de Juan, 18°. Pues siendo Cristo acusado ante Poncio Pilato, vicario del príncipe romano en Judea, por decirse él rey de los judíos, interrogándole Pilato si había dicho tal cosa o se decía él rey, respondió Cristo entre otras cosas a la pregunta de Pilato estas palabras: *Mi reino no es de este mundo*², es decir, no vine a reinar con un régimen o dominio temporal, del modo como reinan los reyes del mundo, en prueba de lo cual en seguida da el mismo Cristo la señal manifiesta: *Si de este mundo fuera mi reino, mis servidores cierto que lucharían para que no fuera entregado a los judíos*. Como si arguyera Cristo de esta manera: Si hubiera venido a reinar en este mundo con gobierno terreno, o sea, coactivo, tendría vasallos de este régimen, luchadores por tanto y debeladores de los trasgresores, como tienen los otros reyes; pero no tengo tales vasallos, como tú manifiestamente puedes echar de ver. De donde la glosa interlinear: *Se ve claro que nadie lo defendió*. y esto es lo que Cristo, reasumiendo por segunda vez, dice: *Ahora bien, mi reino no es de acá, el que efectivamente vine a enseñar*.

[...]

§ 7. De nuevo, para el propósito principal, mirando lo que Cristo mostró con la obra o con el ejemplo. Leemos, en efecto, en Juan, 6°, que *Jesús, cuando se percató de que iban a venir para arrebatarse y hacerle rey, huyó de nuevo al monte él a solas*³. Y la glosa interlinear: *Del cual descendió para dar de comer a las turbas, enseñando a huir de lo próspero de este mundo y a orar en contra de ello*. Ciertamente es, pues, que Cristo huyó del gobierno, y si no, nada nos habría enseñado con el ejemplo. Sentido que refuerza la exposición de Agustín cuando dice que *los fieles cristianos son*

² Juan, 18, 36. *Glossa interl.*

³ Juan, 6, 15.

*su reino, que ahora se cultiva, ahora se compra con la sangre de Cristo. Y se manifestará su reino cuando se revele la claridad de sus santos después del juicio que él hará. y los discípulos y las turbas creyentes en él pensaron que había venido para reinar ya*⁴. Ved cómo los santos por reino de Cristo en este mundo nunca entendieron un dominio temporal o juicio de actos contenciosos y su ejecución por una potestad coactiva contra los trasgresores de las leyes en este mundo, sino cómo por reino de él en el presente siglo entendieron la doctrina de la fe y del régimen acorde con ella con vistas al reino celeste.

[...]

§ 9. Ahora, pues, a continuación, resta mostrar que Cristo mismo, no sólo recusó el principado o el juicio coactivo en este mundo, por lo que dio ejemplo a sus apóstoles y discípulos y a los sucesores de ellos de obrar de la misma manera, sino que mostró con su palabra y con su ejemplo que todos, tanto sacerdotes como no sacerdotes, deben someterse real y personalmente al juicio coactivo de los príncipes de este siglo. Con la palabra, pues, y con su ejemplo mostró esto Cristo, primero, en esas cosas, por lo que tenemos en Mateo, 22°. Pues interrogándole los judíos: *Dinos, qué te parece, ¿es lícito dar el tributo al César o no?*, a éstos Cristo, mirando el denario y su inscripción, les dio por respuesta: *Devolved al César lo que es del César y a Dios lo que es de Dios*⁵. Donde la glosa interlinear: *es decir, el tributo y el dinero*. Y Ambrosio sobre aquello de: *¿de quién es la imagen y la inscripción esa?* dice así: *Como el César exige la impresión de su imagen, así también Dios exige que el alma esté marcada con la luz de su rostro*⁶. Mira, pues, lo que Cristo vino al mundo a exigir. El Crisóstomo dice así: *Pero tú, cuando oigas: devuelve al César lo que es del César, entiende que se refiere sólo a aquello que en nada daña a la piedad, porque si fuera algo de eso, no es ya tributo del César, sino del diablo*⁷. He aquí cómo en todo debemos someternos al César, mientras no esté en contradicción con la piedad, es decir, con el culto o el mandamiento divino. En las cosas reales, pues, quiso Cristo someterse al príncipe secular. Esta fue también claramente la opinión del bienaventurado Ambrosio, apoyándose en la sentencia de Cristo; dijo, en efecto, en la carta contra Valentiniano, que se intitula: *A la plebe: Pagamos al César lo que es del César y a Dios lo que es de Dios. El tributo es del César, no se niega*⁸.

[...]

⁴ *Catena aurea*, XII, 330.

⁵ *Mt.*, 22, 20.

⁶ *Glossa ord.*, ad locum

⁷ *Catena aurea*, XI, 273.

⁸ *Ambrosio, Sermo contra Auxentium de basilicis tradendis*, 35. *PL* 16. 1061

§ 12. Además, no sólo en cuanto a las cosas mostró Cristo que se sometía a la jurisdicción coactiva del príncipe secular, sino también en sí mismo en cuanto a lo personal, mayor que la cual, ninguna jurisdicción, ni sobre él, ni sobre otro pudo tener el príncipe, por lo que también llama a esta jurisdicción el legislador romano *mero imperio*. Esto puede mostrarse con evidencia por Mateo, 27°, porque, como allí se lee y aparece, Cristo permitió ser cogido y conducido al pretorio de Pilato, que era vicario del emperador y por él finalmente, como juez de potestad coactiva, sobrellevó el ser juzgado y entregado al último suplicio, ni contra él, como contra un no juez, se querelló, aunque quizá significó que padecía un juicio menos justo. Y consta que podría haber sufrido tal juicio y tal suplicio de parte de los sacerdotes, si hubiera querido y hubiera juzgado inconveniente que sus futuros sucesores se sometieran a los príncipes de este siglo y fueran juzgados por ellos. Y esto, por estar más de propósito escrito en el 19° de Juan, añadido aquí lo que allí se dice: Cuando, pues, Cristo fue llevado a Pilato, vicario del César como a juez, y acusado de que se hacía rey de los judíos e hijo de Dios, y preguntándole Pilato: *De dónde eres tú*, a lo que Jesús no dio ninguna respuesta, le dijo Pilato las siguientes palabras, muy a punto para nuestro propósito, cuyo pasaje suena así: *¿A mí no me hablas? ¿No sabes que tengo potestad para crucificarte y para soltarte. Respondió Jesús: No tendrías poder alguno sobre mí si no se te hubiese dado de arriba*⁹. He aquí cómo Cristo no negó que Pilato tuviese potestad para juzgar y para ejecutar el juicio sobre él; ni dijo: Esto no te pertenece *de iure* a ti, bien que *de facto* lo haces. Sino que añadió Cristo que esa potestad la tenía Pilato de arriba.

[...]

§ 13. Ni tan sólo de sí quiso Cristo excluir el principado secular o la plenitud judicial coactiva, sino también excluyó de ella a los apóstoles, tanto de ellos entre sí como respecto de los otros. De donde en Mateo, 20°, y en Lucas, 22°, tenemos este pasaje: *Se produjo una gran contienda entre ellos*, es decir, los apóstoles, *sobre quién de ellos era el mayor*. Y les dijo, a saber, Cristo: *Los reyes de las naciones los dominan y los que tienen potestad sobre ellos se denominan bienhechores*. En Mateo se expresa así este inciso: *Y los que son mayores ejercen poder sobre ellos, pero vosotros no así. Sino el que es mayor entre vosotros, hágase como el menor, y el que preside como el servidor. Porque, ¿quién es mayor, el que se recuesta o el que sirve? ¿No es el que se recuesta para comer? Pues yo estoy en medio de vosotros como el que sirve. Pero el que quiera de vosotros ser mayor, sea vuestro servidor. Y el que quiera de vosotros ser el primero,*

⁹ Juan, 19, 9.

sea vuestro siervo, como el hijo del hombre, que no vino para ser servido, sino para servir¹⁰, es decir, servir en las cosas temporales, no dominar o anteponerse. Porque en el ministerio espiritual era él el primero, no siervo, en medio de los apóstoles. Donde Orígenes: *Sabéis que los príncipes de los gentiles los dominan, es decir, no contentos sólo con gobernar a sus súbditos, tratan de dominarlos violentamente*, a saber, por el poder coactivo, si hace falta. *Mas entre vosotros, que sois míos, no será así. Porque como todo lo carnal está fundado en la necesidad, y lo espiritual en la voluntad, así los que son príncipes espirituales, es decir, los prelados, han de fundar su supremacía en el amor, no en el temor*¹¹. El Crisóstomo por su parte, entre otras cosas, dice esto que hace a nuestro propósito: *Los príncipes de este mundo están para dominar a sus inferiores y reducirlos a esclavitud y despojarlos, suple, si desmerecen, y hasta la muerte utilizarlos para su provecho y gloria, se entiende, del principado. Y los príncipes de la iglesia, es decir, los prelados, son hechos para que sirvan a sus inferiores y les suministren todo cuanto recibieron de Cristo, descuiden su propio provecho y procuren el de aquéllos y no retrocedan ante la muerte por la salvación de sus inferiores. Y codiciar el primado de la iglesia no es injusto ni provechoso. Porque, ¿qué hombre sensato quiere por propia iniciativa someterse a la servidumbre y a un peligro tal que tenga que responder de toda la iglesia, si no es acaso el que no teme el juicio de Dios, aprovechándose de su primacía eclesiástica de modo asegurado hasta convertirla en secular?*¹² Pues, ¿para qué tienen que entrometerse los sacerdotes en los juicios seculares? Porque no deben dominar temporalmente, sino, a ejemplo y mandato de Cristo, servir. De donde Jerónimo: *Finalmente les propone su ejemplo, se entiende, Cristo, de modo que si menospreciaren las palabras, suple, del mismo, los apóstoles se avergüencen de los hechos, de dominar temporalmente*¹³. Y Orígenes parecidamente, sobre aquello: *Y dar la vida por la redención de muchos*¹⁴, dice así: *Los príncipes, pues, de las iglesias deben imitar a Cristo accesible y hablando a las mujeres e imponiendo las manos a los niños, lavando los pies de los discípulos, para que lo mismo hagan ellos con los hermanos. Pero nosotros somos tales, habla de los prelados de su tiempo, que no parece sino que aventajamos en la soberbia aun a los príncipes del mundo, o no entendiendo o menospreciando el mandato de Cristo y ambicionamos, como los reyes, tener escuadrones que marchen delante y terribles*¹⁵. Pues

¹⁰ Lc., 22, 24-27; Mt., 20, 25-28.

¹¹ *Catena aurea*, XI, 234.

¹² *Catena aurea*, *ibíd.*

¹³ *Catena aurea*, *ibíd.*

¹⁴ Mt., 20, 28.

¹⁵ *Catena aurea*, *ibíd.* Lc., 22, 24-27.

porque hacer esto es menospreciar o ignorar el mandato de Cristo, se ha de amonestar primero a estos prelados, lo cual haremos en este tratado, mostrando qué es lo que les compete por su autoridad; después, si tomaran a menos el corregirse, se les ha de forzar y obligar por los príncipes seculares, para que no corrompan las costumbres de los otros.

[...]

Así pues, de las verdades evangélicas aducidas, y de sus interpretaciones dadas por los santos y por otros doctores aprobados, se hará a todos evidente que Cristo se excluyó a sí mismo o se quiso excluir, ya de palabra ya de obra, de todo principado o gobierno, juicio o potestad coactiva terrena, y que quiso someterse a los príncipes y potestades del siglo con jurisdicción coactiva.

V. DE LOS DICHOS CANÓNICOS DE LOS APÓSTOLES Y DE LAS EXPOSICIONES
DE LOS SANTOS Y DE LOS DOCTORES, POR LAS QUE ABIERTAMENTE SE PRUEBA
LO MISMO QUE EN EL CAPÍTULO PRECEDENTE

Resta, pues, y sobre esto se ha de mostrar, que la misma opinión y doctrina fue de los principales apóstoles, primero de Pablo en la 2ª a Timoteo, 2º, el mismo que él había constituido como presbítero u obispo, amonestándole a que no se mezclara en los negocios humanos. Dice él así: *Nadie que milita para Dios se enreda en negocios seculares*¹. Donde la glosa, según Ambrosio: *Porque nadie que milita en las cosas espirituales para Dios, que no puede dividirse en dos siervos contrarios, como nadie puede servir a dos señores, se mezcla en cualesquiera negocios seculares*², y dijo *cualquiera*, sin excluir ninguno. Y dado que es secularísimo entre los negocios seculares el principado o el juicio coactivo de los actos contentiosos, pues ordena y regula todos los negocios seculares o actos seculares humanos civiles, como se demostró en el 15º de la Primera Parte, principalísimamente mandó el apóstol huir de él al que debe militar para Dios, a saber, administrando las cosas espirituales, cual debe ser el obispo y cualquier sacerdote.

§ 2. Y ésta, que dijimos ser, la opinión del apóstol, más abiertamente la expresa el pasaje de la 1ª a los Corintios, 6º, cuando dice: *Si tuviereis negocios seculares, echad mano, para juzgarlos, de los menos considerados dentro de la iglesia*³. Pues allí hablaba el apóstol a todos los

¹ II Tim., 2, 4.

² Collect., PL 192, 1577 B-C. Marsilio atribuye el texto a Ambrosio.

³ I Cor., 6, 4.

fieles y a la iglesia propiamente dicha, a saber, según la última significación.

[...]

Y esto lo expresa de nuevo cuando añade: *A mí no me parece buen apreciador de las cosas quien juzga indigno de los apóstoles o de los varones apostólicos el no erigirse en juez de esas cosas, a los que se les dio un oficio para cosas mayores. ¿Cómo no van a tener en poco el juzgar de hacendillas terrenas de los hombres los que en el reino celeste juzgarán aun a los ángeles?*⁴

§ 4. Mandó también el santo apóstol a todos, sin exceptuar a ninguno, que el obispo o el sacerdote o diácono estuvieran sometidos al juicio coactivo de los jueces y príncipes seculares, y no resistirles, a no ser que mandasen hacer algo contra la ley de la salud eterna. De donde a los Romanos, 13º: *Toda alma esté sometida a las potestades superiores; porque no hay autoridad sino venida de Dios. Y las cosas que vienen de Dios son ordenadas. Así que quien resiste al poder resiste a la ordenación de Dios. Los que resisten se acarrearán a sí mismos la condena. Porque los príncipes no están para temor de la obra buena, sino de la mala. ¿Quieres no temer a la potestad? Obra el bien, y tendrás aplauso de ella, pues es para ti ministro de Dios para el bien. Mas si obras mal, teme, porque no sin razón lleva la espada. Ministro es de Dios, vengador en su enojo para aquel que obra el mal. Por eso estad sumisos, no sólo por temor de la ira, sino también por razón de la conciencia. Por eso también pagáis tributos, porque son ministros de Dios, que sirven a Dios en eso mismo. Dad a todos lo debido, al que tributo, tributo, al que impuesto, impuesto, al que temor, temor, al que honor, honor*⁵. Quise aducir las glosas sobre estas notables palabras del apóstol, según los santos y doctores católicos, porque de lo que dijo el apóstol y los glosadores tan abiertamente se demuestra que nuestro propósito va en la verdad, que nadie de sanamente puede, bien mirado eso, dudar ya. Dijo, pues, el apóstol: *Toda alma*, etc., sin exceptuar a nadie, donde la glosa según Agustín en primer lugar y luego a veces según Ambrosio así: *Y aquí invita a la humildad. Porque a algunos parecía que los malos amos y, sobre todo, los infieles, no debían ejercer su dominio sobre los fieles, y si fueren buenos y fieles, debían ponerse en línea de igualdad con los buenos y fieles. Soberbia ésta que también reprueba el apóstol significando por la parte superior, es decir, el alma, todo el hombre. Porque, ¿qué es toda alma sino todo hombre? Como si dijera: “Todo*

⁴ Bern., *De consid.*, I, 6-7. PL 182, 735 C-736 A.

⁵ Rom., 13, 1-7.

lo antedicho hay que hacerlo; aunque seáis tan perfectos en el cuerpo de Cristo, sin embargo, toda alma ha de someterse, es decir, sométase todo el hombre. Al cual significó por el nombre de alma, a fin de que no sirváis sólo con el cuerpo, sino también con la voluntad". Toda alma, pues, esté sometida, de modo que también de voluntad sirva el hombre a las potestades seculares, buenas o malas, a saber, a los reyes, a los príncipes, a los tribunos, a los centuriones y otros así⁶. He ahí, pues, lo que entendió el apóstol por las potestades superiores; no otra cosa que los príncipes seculares.

[...]

§ 5. De este pasaje del apóstol y de las explicaciones aducidas de los santos, se deduce que todo aquel que no quiera que se blasfeme el nombre y las enseñanzas del Señor como si fueran injustas, predicando contra las leyes civiles, como dice la glosa de Agustín a la 1ª a Timoteo, 6º, sin dubitación alguna ha de sostener que todos los hombres, de cualquier estado y condición que sean, real y personalmente deben sujetarse a la jurisdicción de los príncipes de este mundo y obedecerles en todo aquello que no contradiga a la ley de la salud eterna, máxime ateniéndose a las leyes humanas o a las costumbres honestas y aprobadas; porque de éstas habla claramente el apóstol cuando dice: *Toda alma se someta*, etc., y que *no sin motivo llevan la espada*, y lo demás que sobre los mismos expresó, aludiendo tanto a la defensa que hacen de la patria, como a la prestación a ellos de los tributos, según las explicaciones de los santos.

[...]

§ 6. No queremos, sin embargo, decir con todo esto que no se deba reverencia y obediencia al doctor o pastor eclesiástico en aquellas cosas que manda o enseña dentro y según la ley evangélica, no de otra manera o contrariamente a ella, como por Mateo, 23º, y por Jerónimo glosándolo aparece bastantemente. Bien que tampoco debe ni puede él en este mundo forzar a nadie a la observancia de aquellas cosas bajo pena o suplicio ninguno, real o personal, puesto que no leemos en la escritura evangélica que tal potestad de forzar o de dominar haya sido concedida a nadie en este mundo, sino más bien denegada de consejo o de precepto, como de lo dicho aquí y en el capítulo precedente se desprende. Pues tal potestad en este siglo se da por las leyes o por los legisladores humanos, la cual, dada al obispo o al presbítero para forzar a los hombres en aquellas cosas que pertenecen a la ley divina, sería inútil. Pues a los forzados así en nada les aprovecharía para la salud eterna.

[...]

⁶ *Collect.*, PL 191, 1503 A-B.

[P]orque, en efecto, no está en su poder en forzar a nadie, porque no tienen en este siglo autoridad coactiva sobre ninguno, por razón de que *de las leyes*, es decir, de los legisladores, no les vino a ellos aquella autoridad en aquellos tiempos o en aquellos lugares o provincias. Entonces, apelando a otra razón, añade: *Ni si fuera dada*, se entiende, una tal potestad, *tendríamos* nosotros, es decir, los obispos o presbíteros, *donde ejercitar aquella potestad, dado que nuestro Dios*, a saber, Cristo, *va a remunerar, no a los que son apartados a la fuerza del pecado*, es decir, ejerciendo sobre ellos violencia, *sino a los que se guardan de él por propia voluntad*⁷.

§ 7. No queremos, sin embargo, decir con esto que sea desaconsejado ejercer coacción sobre los herejes y otros infieles, pero aseveramos que esta autoridad, si ello entra en lo lícito, compete exclusivamente al legislador humano.

No le compete, pues, a cualquier sacerdote u obispo la potestad coactiva, sino deben, igual ellos que los demás, someterse en esto a los jueces seculares en esa potestad, como hemos dicho.

[...]

§ 10. No habiendo sido, pues, concedida por ley divina a ningún obispo la jurisdicción o potestad coactiva en este mundo, sino más bien denegada como consejo o como precepto, como meridianamente se ha demostrado en este capítulo y en el precedente, y no perteneciendo tampoco tal potestad a los obispos o presbíteros, en cuanto obispos por sucesión hereditaria o paterna, síguese necesariamente que en esa potestad están ellos sujetos a los jueces seculares, como se deduce con evidencia de los dichos de los apóstoles Pedro y Pablo y de los otros santos, y con discurso demostrativo puede hacerse patente por lo dicho en los capítulos XV y XVII de la Primera Parte, y que ni obispo alguno ni el Papa tienen jurisdicción coactiva sobre ningún presbítero o no presbítero, en este siglo, a no ser que les sea concedida por el humano legislador, en cuya potestad está siempre el retirársela por motivos razonables, cuya apreciación y determinación plena se entiende pertenecer al mismo legislador, sobre todo en las comunidades de fieles.

Que, pues, Cristo renunció a cualquier principado o jurisdicción coactiva sobre alguien en este mundo, y que se la quitó a sus apóstoles y sus sucesores los obispos o presbíteros, o por consejo o por precepto, y que a sí mismo y a los mismos apóstoles los quiso sometidos a la jurisdicción coactiva de los príncipes del siglo, y que enseñó a guardar lo mismo, y como él

⁷ I Crisóst., *De sacerdocio*, II, 3; PG 48, 634.

los principales apóstoles, Pedro y Pablo, de obra y de palabra; por las verdades del evangelio, por testimonios de valor eterno y por las interpretaciones y exposiciones de los santos y de otros probados doctores de la fe cristiana, creemos haberlo demostrado con evidencia.

[...]

VII. PRESENTACIÓN SUMARIA DE LO DICHO EN EL CAPÍTULO PRECEDENTE Y DE SU ACLARACIÓN Y CONFIRMACIÓN

Pues tratando de poner en resumen lo que toca a la potestad de las llaves sacerdotales concedida por Cristo a los apóstoles, diremos que en el pecador verdaderamente penitente, es decir, dolido del pecado cometido, unas cosas obra Dios solo, sin previo ministerio alguno del sacerdote, a saber, la iluminación de la mente, la purificación de la culpa o mancha del pecado, y la remisión de la condenación eterna. Otras son las que Dios obra en el mismo pecador, no por sí solo, sino mediante el ministerio del sacerdote, como mostrar ante la faz de la iglesia quién se ha de tener por desligado o ligado de sus pecados en este mundo, quién ligado o desligado en el otro; es decir, a quién Dios ha remitido los pecados o a quiénes se los ha retenido. Aún hay algo que Dios obra sobre el pecador por el ministerio del sacerdote, a saber, la conmutación de la pena del purgatorio merecida por el pecador para el otro mundo a cambio de alguna satisfacción temporal en éste. Porque remite de aquella pena en parte o en todo, según las satisfacciones impuestas y según la disposición del pecador. Todo lo cual ha de ser hecho por el sacerdote con la llave de la potestad, según su discreción. Y así también han de ser excluidos de la comunión de los sacramentos por el sacerdote los contumaces y admitidos los arrepentidos, como se dijo hacia el fin del capítulo precedente.

§ 2. Y ésta fue la opinión del Maestro, libro 4º, Distinción 18ª, cap. 8º, cuando dice: *Según estos modos de atar y desatar, en qué sentido es verdadero lo que se dice: “Todo lo que desates sobre la tierra, desatado será en el cielo y todo lo que ates en la tierra atado será en el cielo”.* Porque a veces dan, se entiende, los presbíteros, por desatados o atados los que no lo son ante Dios, y alguna vez atan o desatan con la pena de la satisfacción a los indignos, y admiten a los sacramentos a los indignos y se retraen de admitir a los dignos. Pero hay que entenderlo de aquéllos cuyos méritos piden ser desatados o atados. Así pues, todos aquéllos a los que desatan o atan con la llave de la discreción, aplicada a los méritos de los reos, son desatados o atados en los cielos, es decir, ante Dios, pues entonces la sentencia proferida por el sacerdote es aprobada y confirmada por

*el juicio divino. Y dice el Maestro a seguido, a modo de epílogo: He aquí cuál es y hasta dónde se extiende el uso de las llaves apostólicas*¹.

§ 3. Y para tener mayor evidencia de ello, aducimos el ejemplo o comparación familiar a él, que parece tener “también mucho que ver con las palabras y la sentencia de Cristo y de los santos, cuyas autoridades mencionamos antes, máxime la de Ambrosio. Dice, pues, aquel que *la palabra de Dios perdona el pecado y el sacerdote es el juez; el sacerdote desempeña su oficio, pero no ejerce los derechos de potestad alguna*². Entonces, ¿cómo perdona el pecado el sacerdote si no ejerce los derechos de potestad alguna? Digamos, pues, que el sacerdote, como llavero del juez celeste, libera al pecador de un modo análogo al del llavero del juez terreno. Porque como por la palabra o sentencia del juez de este mundo, es decir, del gobernante, se le condena o se le absuelve de la culpa y de la pena civil al reo, así por la palabra divina se absuelve, absolutamente a alguno o se le ata con la culpa y la deuda de la condenación eterna o pena para el siglo venidero. Y como por la acción del llavero del príncipe terreno nadie es absuelto o condenado a la culpa y pena civil, pero por la acción de cerrar o abrir la cárcel queda a la vista que el reo es absuelto o condenado, de modo parecido por la acción del sacerdote ninguno es desatado o atado con la culpa y la deuda de la condenación eterna, sino que se muestra ante la faz de la iglesia quién se ha de tener por absuelto o atado por Dios, al tiempo que recibe la bendición del sacerdote y se le admite a la comunión de los sacramentos, como dejamos dicho al final del capítulo precedente. Y por eso, como el llavero del juez terreno abriendo y cerrando la cárcel cumple su oficio, sin que por ello ejerza derechos algunos de potestad judicial de absolver o de condenar, porque, aunque de hecho abra la cárcel a algún reo no absuelto por el juez y muestre al pueblo con su voz que el reo ha sido absuelto, no por ello el reo aquel quedaría absuelto de la pena o de la culpa civil, y, al revés, si rehusara abrir la cárcel y con su voz proclamará que no ha sido absuelto sino condenado aquél al que verdaderamente el juez hubiera absuelto por sentencia suya, no por ello sería deudor de la culpa o pena civil, así análogamente el sacerdote, llavero del juez celeste, cumple con su oficio pronunciando verbalmente la absolución, o la atadura, o condenación. Que si a los que verdaderamente el juez de arriba habrá de condenar o ha condenado ya, el sacerdote, por ignorancia o por móviles subrepticios, o por ambas cosas, proclamare dignos de absolución o ya absueltos, o, al revés, no por ello serían los primeros absueltos y los segundos condenados,

¹ P. Lomb., *Sent.*, IV, *Dist.* 18, c. 8. *PL* 192, 888.

² *Ibíd.*

por no haber hecho uso de la llave o de las llaves con discreción, según los méritos de los reos.

[...]

§ 5. Por lo cual el sacerdote conforme a su oficio no ha de ser comparado similarmente con el juez en la tercera significación, sino en la primera, a saber, con el que tiene autoridad para enseñar y obrar a la manera del médico, pero no con potestad coactiva sobre ninguno. El mismo Cristo se llamó a sí mismo médico en este sentido, no príncipe, como en Lucas, 5°, dijo hablando de sí: *Que no tienen necesidad de médico los sanos sino los enfermos*³. No dijo: necesitan de juez. Porque no había venido a este mundo para ejercer el juicio coactivo en los litigios contenciosos humanos, como lo declaramos según Lucas, 12°, en el cap. IV, párrafo 8. A los vivos y a los muertos los juzgará con tal juicio coactivo en aquel día del que habla el apóstol, en la 2ª a Timoteo, cap. último: *Por lo demás me está reservada a mí la corona de la justicia, que me será dada en aquel día por el Señor, como justo juez*⁴. Entonces, en efecto, impondrá con juicio coactivo las penas a aquellos que hubieren trasgredido su ley dada inmediatamente por él mismo en este mundo. Y por ello significativamente dijo Cristo a Pedro: *A ti te daré las llaves del reino de los cielos*. Y no dijo: A ti te daré el juicio del reino de los cielos. Y por lo mismo, ya dijimos que el llavero del juez mundano o del excelso juez no tiene el juicio coactivo, el que denominamos como tercera significación, por razón de que ni uno ni otro ejerce los derechos de tal potestad, como claramente lo dijo Ambrosio del sacerdote y también sobradamente se ha demostrado con la autoridad de los otros santos.

Así, pues, sobre la autoridad de los sacerdotes u obispos y de la potestad de las llaves apostólicas, por Cristo a ellos concedida, quede en firme lo dicho.

[...]

XV. DE LA DISTINCIÓN DEL OFICIO SACERDOTAL, SEGÚN SU AUTORIDAD ESENCIAL Y ACCIDENTAL, SEPARABLE E INSEPARABLE, Y QUE EN LA DIGNIDAD ESENCIAL NO ES INFERIOR CUALQUIER PRESBITERO AL OBISPO, SINO SOLO EN LA ACCIDENTAL

Junto con todo lo dicho se suscitó una cuestión difícil y muy necesitada de someterse a consideración. Pues se dijo por nosotros en el XV de la Primera Parte y se volvió a tocar en algún modo en el XVIII, al fin, que el legislador humano mismo, o por sí, o por medio de la parte gobernante, es

³ Lc., 5,31.

⁴ II Tim., 4, 8.

la causa eficiente de la institución de todos los oficios o partes de la ciudad. Con esto recordamos también haber dicho en el último de la Primera que el sacerdocio o el oficio sacerdotal de la Nueva Ley fue en un principio instituido por Cristo solo, del cual mostramos que había abdicado todo principado de este siglo y todo el dominio de las cosas temporales, en el IV, XI, XIII y XIV de esta Parte, y que no fue legislador humano, en el XII y XIII de la Primera. Por lo cual creemos haber afirmado que el mismo no fue el instaurador de oficio alguno en la ciudad, ni humano legislador o príncipe, de donde con razón se preguntará alguien: ¿A quién le corresponde la autoridad de instituir el sacerdocio, máxime en las comunidades de fieles? Porque lo dicho por nosotros sobre ello parece contradecirse en sus términos.

§ 2. Proponiéndonos solventar la aparente contrariedad de lo dicho, traeremos primero a la memoria lo que dijimos en el VI y VII de la Primera Parte, que unas son las causas de cualquiera de los oficios de la ciudad, según que los oficios designan hábitos del alma, y otras las causas de los mismos según que son partes de la ciudad, instituidos para obtener de ellos las correspondientes suficiencias, lo cual puede análogamente mirarse en el sacerdocio, como en los demás oficios de la ciudad. Porque del sacerdocio, según que designa un hábito del alma, que llaman carácter los doctores de las sagradas escrituras, la causa eficiente inmediata por sí, o el propio hacedor de suyo, es Dios, quien imprime ese carácter en el alma, previo un cierto ministerio humano que lo prepara. El comienzo de ello en la Ley Nueva provino de Cristo. Porque él, que fue verdadero Dios y verdadero hombre, en cuanto sacerdote humano desempeñó un ministerio que, siguiéndole, lo practican ahora los sacerdotes; mas en cuanto Dios imprimió un carácter en las almas de aquellos que instituyó sacerdotes; del cual modo, primero instituyó a los apóstoles, como sucesores suyos inmediatos, y así después, sucesivamente, a todos los demás sacerdotes, pero por el ministerio de los apóstoles y de todos los demás que van sucediéndose en este oficio. Porque imponiendo los apóstoles y los otros sacerdotes las manos a otros, y profiriendo las correspondientes palabras y oraciones prescritas para ello, Cristo, en cuanto Dios, imprime ese hábito o carácter sacerdotal en los que dignamente quieren recibirlo.

Y semejantemente hay que pensar de las colaciones de los otros órdenes por los que se imprime un cierto carácter en el alma del que los recibe. Este carácter sacerdotal, pues, ya sea uno, ya varios, es una potestad por la que el sacerdote puede consagrar del pan y del vino el Cuerpo y la Sangre bendita de Cristo, aplicando la pronunciación de ciertas palabras, y

administrar todos los demás sacramentos eclesiásticos; por la cual puede igualmente desatar o atar a los hombres sus pecados.

[...]

§ 4. Este carácter sacerdotal, sea uno o muchos, el que hemos dicho ser potestad de realizar el sacramento de la Eucaristía, o del Cuerpo y Sangre de Cristo, y potestad también de desatar o atar a los hombres de sus pecados, y al que en adelante llamaremos autoridad esencial, o sea, la que es inseparable del presbítero en cuanto presbítero, me parece con probabilidad que todos los sacerdotes la poseen igual en la especie, y que ni el obispo romano ni otro cualquiera la tienen en mayor grado que cualquier así dicho simple sacerdote. Pues en esta autoridad, sea una o muchas, el obispo no difiere del sacerdote, según el testimonio de Jerónimo, o con más verdad del mismo apóstol, de quien es manifiesta la sentencia, como más abajo se verá claro. Dice, pues, Jerónimo sobre aquello de Mateo, 16°: *Y cualquier cosa que atares sobre la tierra, etc.: Tienen la misma potestad judiciaria los otros apóstoles*, suple, que tuvo Pedro, *a los que, después de la resurrección, dijo*, se entiende, Cristo: *Recibid el Espíritu Santo; y a quienes perdonareis los pecados les serán perdonados*, etc. *Lo tiene toda la iglesia en los presbíteros y en los obispos*¹; antepone aquí a los presbíteros, porque la autoridad ésta se le debe al presbítero, como presbítero que es, primariamente y en cuanto tal. De la potestad referida al sacramento de la Eucaristía, nadie discute que sea igual en cualquier sacerdote que en el romano pontífice. Y por ello es de admirar cómo algunos² se empeñan pertinazmente en sostener, pero con menos razón, que el romano pontífice tiene de Cristo una más amplia potestad de llaves que los demás sacerdotes, cuando por la Escritura esto no se demuestra, sino más bien lo opuesto.

[...]

§ 6. Pero quedando atrás los tiempos de los apóstoles y creciendo notablemente el número de los sacerdotes, para evitar el escándalo y el cisma, eligieron los sacerdotes uno entre ellos que dirigiera a los otros y ordenara cuanto concierne a ejercer el oficio y servicio eclesiástico, como repartir las ofrendas y disponer todo lo demás del modo más conveniente, no fuera que, obrando cada cual a su talante y a veces no debidamente, se turbara la organización y el servicio de los templos por la diversidad de apetencias. Este elegido, pues, para regular a los demás sacerdotes, por la costumbre de los posteriores, retuvo para sí solo el nombre de obispo, como superintendente, porque no sólo era superintendente sobre el pueblo

¹ Mt., 16, 19. *Glossa ord., ad locum.*

² Alusión probable a Bern., *De consid.* II, 8; PL 182, 751-752 y a Tomás de A., *Summa Theol., Supplem.*, q. 24 a. 1., ad l.

fiel, por lo que todos los sacerdotes de la iglesia primitiva se denominaban obispos, sino porque también era superintendente sobre los otros co-presbíteros suyos, por lo cual uno de Antioquía retuvo para sí solo el nombre de obispo, reteniendo los demás para sí el simple nombre de sacerdotes.

[...]

§ 9. Existen otra u otras diferentes instituciones no esenciales de los oficios sacerdotales, como la que hemos ya llamado una determinada elección, por la cual uno de entre ellos es asumido para la ordenación y dirección de los otros para todo lo tocante al culto divino; como son las elecciones e instituciones de algunos para enseñar e instruir, y para administrar los sacramentos de la Nueva Ley a un cierto pueblo o determinado lugar, mayor o menor, igualmente para distribuir, tanto entre sí, como entre los demás pobres, ciertos bienes temporales establecidos y ordenados, por el legislador o por personas particulares, para sustentación de los evangelizadores pobres en determinada provincia o comunidad, y también para la sustentación de otros pobres que no pueden procurarse lo necesario para sí o por la edad o por la enfermedad, o por otra causa lamentable, bien entendido, de aquello que sobrare de lo destinado para la suficiencia de los evangelizadores. Estos bienes temporales, así establecidos, se llaman, con palabra moderna, *beneficios eclesiásticos*. [...].

§ 10. Pues así de lo dicho se saca de quién sea, como su causa eficiente, la institución del sacerdocio, en cuanto que significa un hábito del alma, como un carácter suyo, y de los otros órdenes que llaman sagrados, a saber, que han sido instituidos inmediatamente por Dios o por Cristo, aunque mediante un cierto ministerio humano preparatorio, como la imposición de las manos y la pronunciación de palabras, que acaso nada hacen respecto de esto, pero que así se hacen preceder por un cierto pacto u ordenación divina.

Aparece también por lo dicho que en la autoridad primera, que en un principio llamamos esencial, todos los sacerdotes son iguales en mérito y en el mismo sacerdocio como lo dijo Jerónimo en la carta citada, asignando la causa, porque *todos los obispos son sucesores de los apóstoles*. En lo cual parece insinuar que todos los apóstoles fueron de igual autoridad; y así, por consiguiente, ninguno de ellos tuvo personalmente autoridad sobre otro, o sobre los demás, ni en cuanto a la institución esencial ni en cuanto a las instituciones secundarias.

De donde parece que igualmente hay que opinar sobre los sucesores de ellos entre sí, para los cuales y sobre los cuales nos queda determinar de dónde vengan las instituciones que hemos llamado secundarias y creadas por la autoridad humana, y cuál sea razonablemente la causa eficiente de ellas.

XVI. DE LA IGUALDAD DE LOS APÓSTOLES EN EL OFICIO O CUALQUIER DIGNIDAD
CONFERIDA A ELLOS INMEDIATAMENTE POR CRISTO. DE DONDE SE PRUEBA LO DICHO
EN EL CAPÍTULO PRECEDENTE SOBRE LA IGUALDAD DE TODOS LOS SUCESORES
Y EN QUÉ SENTIDO TODOS LOS OBISPOS SON INDIFERENTES SUCESORES
DE UN APÓSTOL CUALQUIERA

[...]

§ 19. Y lo que sobre todo y como lo principal hay que tener presente, por lo que toca a nuestro propósito, que si existen algunas congruencias, por las cuales aparezcan los sucesores del bienaventurado Pedro en algún modo más venerables que los demás, y principalmente en la sede episcopal, ninguna necesidad ni voz hay allí de la Sagrada Escritura por la que los sucesores de los otros apóstoles se crean sujetos a ellos por alguna susodicha potestad. Esto aunque demos que la autoridad de los apóstoles no fuera igual en todos. No tuvo, pues, el bienaventurado Pedro ni ningún otro apóstol, en virtud de las palabras de la Escritura, potestad de instituir o deponer a los otros, ni en lo tocante a la dignidad sacerdotal, la que dijimos esencial, ni en lo tocante a la misión de ellos, o designación para un lugar o pueblo, ni en lo tocante a la interpretación de la Escritura o de la fe católica, ni en cuanto a la jurisdicción coactiva sobre nadie en este siglo, más que lo opuesto. De donde necesariamente parece seguirse que, en virtud de las palabras de la Escritura, ningún sucesor de alguno de ellos, como quiera que sea singularmente llamado, tiene potestad alguna de las dichas sobre los sucesores de los otros. Puede esto patentizarse con la autoridad de Jerónimo, que adujimos de la carta de él a Evandro, al final del párrafo 8º del capítulo precedente.

XVII. DE LA AUTORIDAD DE INSTITUIR OBISPOS Y OTROS CURATOS
Y LOS DEMÁS MINISTROS DE LA IGLESIA EN LO TOCANTE A LA DOBLE DIGNIDAD
U OFICIO, SEPARABLE O INSEPARABLE

Dejamos así ya dicho sobre la causa eficiente de la primera autoridad sacerdotal, que hemos denominado esencial. De la otra, a saber, por la que al frente de ciertos sacerdotes o pueblo, o al frente de ambos son colocados algunos sacerdotes en determinada provincia o lugar; también, de dónde venga el poder de dispensar los bienes temporales llamados *beneficios eclesiásticos*; de dónde igualmente les vengan a ellos o a algunos de ellos las jurisdicciones coactivas; a quién o a quiénes además, y de qué modo más conveniente, pertenezca el determinar los sentidos dudosos de la Sagrada Escritura, máxime de los que tocan a la necesidad de la salud, es lo

que nos queda por definir. Porque suficientemente declarados estos puntos, aparecerá lo que teníamos propósito de poner de manifiesto en el principio de esta obra.

§ 2. Mas antes que abordemos en detalle esto que nos proponemos, conviene narrar primero el modo de institución y determinación de los obispos o presbíteros, según el estado y el comienzo de la iglesia primitiva, de donde después todo lo demás ha derivado. El principio de todo esto hay que tomarlo de Cristo, que es la cabeza y la piedra sobre la cual está fundada la iglesia católica, conforme a lo que dice el apóstol a los Efesios, 4° y 5°¹ y 1ª a los Corintios, 10°². Lo cual dice también cierta glosa sobre aquello de Mateo, 16°: *Sobre esta piedra edificaré mi iglesia*³; digo que esta cabeza de la iglesia, piedra y fundamento, a saber, Cristo, confirió a los apóstoles, a cada uno y por separado, el sacerdocio y la dignidad episcopal sobre todas las naciones y pueblos, no limitando a ninguno para determinado lugar o pueblo, de manera que no le fuera lícito también predicar en cualquier parte, aunque algunos estuvieran más deputados para los gentiles, otros para la circuncisión, por ordenación entre sí y por guía del Espíritu Santo.

[...]

§ 5. Esto así puesto por delante, entrando más en detalle en lo que nos proponíamos al principio de este capítulo, mostraremos, lo primero, que la causa eficiente más conveniente inmediata de la determinada institución de los apóstoles para ciertos pueblos y provincias fue la revelación de Cristo o la ordenación concorde de ellos entre sí; después que la causa eficiente inmediata de una determinada institución de sus primeros sucesores antes de la conversión del pueblo, fue la expresa voluntad de todos los apóstoles o de muchos, si estaban presentes todos o muchos de ellos en la provincia en la que había que establecer un sacerdote o un obispo, o de uno solo de los mismos apóstoles según las circunstancias de lugar, del pueblo o del tiempo; finalmente, que tras la muerte de los apóstoles o en ausencia de ellos, la institución secundaria de los obispos y de los otros ministros de la iglesia o ministros espirituales, se llevó a cabo del modo más conveniente de todos los posibles, mirando a la convivencia humana, por la totalidad de los fieles en el lugar o provincia, sobre lo cual los dichos ministros han de instituirse, nunca por algún otro colegio o grupo particular, ni por una persona. Y así después lo demás se hará claro consecuentemente.

¹ Efes., 4, 15; 5, 23.

² I Cor., 10, 4.

³ Mt., 16, 18. *Glossa or., ad locum.*

De esto, pues, se puede mostrar, lo primero, que no se puede asignar una más conveniente causa de aquella determinación que la divina revelación o la común deliberación de aquéllos, porque en ninguna de ellas parece intervenir error o malicia. De la divina revelación, ninguna duda; de la elección de los apóstoles, parece probable y creíble, pues estaban inspirados por el Espíritu Santo, como ya adujimos de Juan, 20^{o4}.

§ 6. Después de esto digo que la causa eficiente inmediata de la institución o determinación de sus primeros sucesores, máxime antes de la conversión de los pueblos, fueron y tuvieron que ser todos o los más de los apóstoles o uno solo de ellos del modo predicho, si encontraban allí al mismo tiempo, todos, o muchos o sólo uno. Esto se demuestra, primero por la Escritura, pues leemos por los Hechos, 6^o, que así se condujeron en la institución de los diáconos, aun en cuanto a la primera autoridad, la que denominamos esencial. De donde: *Pusieron a éstos ante la mirada de los apóstoles, y orando*, se entiende, los apóstoles, *les impusieron las manos*⁵. No los llevaron, pues, a Pedro solo, sino *ante la mirada de los apóstoles*. Ni se tomó para sí solo Pedro la autoridad de imponerles las manos, sino que los apóstoles *les impusieron las manos*.

[...]

§ 7. Pero si no estuviesen todos o la mayor parte juntos para poner un obispo al frente de la multitud de fieles, para custodiarlos y conservarlos en la fe, hay que decir que esto podía hacerlo uno sólo lícitamente, máxime tratándose de una comunidad pequeña de fieles, ruda e inexperta para discernir qué persona más idónea se encontraba allí para el oficio del episcopado, y sobre todo donde fuera prácticamente nula la suficiencia de personas idóneas para dicho oficio, tal como le aconteció muchas veces a Pablo y a sus primeros sucesores por las causas predichas, como bien aparece por los Hechos y por sus cartas a Timoteo y a Tito. Y que tal institución haya podido y debido realizarse por uno solo lícitamente, puede mostrarse, porque de este modo era elegido el mejor y más conveniente pastor. Porque, o era lícito preferirse uno a sí mismo a otros, en orden al ministerio del evangelio, por su propia voluntad, o esto debía hacerse por la elección de la multitud llana, o por alguno de los apóstoles allí presente. Del primer modo podía seguirse escándalo y error; escándalo, en efecto, si dos o tres hubieran querido asumir para sí aquella autoridad; error o insuficiencia del pastor semejantemente, porque por la mayor parte los necios o los ambiciosos anhelan la presidencia y tratan de alcanzarla más que los varones virtuosos y sabios. Del segundo modo, a saber, que por la elección de aquella multitud sea constituido alguien en prelado, ocurriría verosíblemente el error y la

⁴ Juan, 20, 22-23.

⁵ Hechos, 6, 6.

incompetencia, por la inmadurez de la multitud aquella en cuanto tal, en cantidad y en calidad; porque incultos y manejables abundaban al principio en muchas provincias, principalmente fuera de Judea, como se deduce de la carta a los Gálatas y de muchas otras. De donde el apóstol en la 1ª a los Corintios, 3º: *Y yo, hermanos, no pude hablarlos a vosotros como a espirituales, sino como a carnales. Como a párvulos en Cristo os di a beber leche. Porque aún no lo podíais, ni ahora lo podéis, pues aún sois carnales*⁶. Por ello más seguro y más saludable hacerse esta institución por la elección o determinación de alguno de los apóstoles, cuya vida y sabiduría era preponderante y más amplia, por el hábito del Espíritu Santo, más queja de todos juntos en la tal multitud, si bien no se ha de negar que fue conveniente que el apóstol consultase a la multitud sobre las costumbres del que iba a ser puesto al frente.

§ 8. Después de esto quiero demostrar que, después de los tiempos de los apóstoles y de los primeros padres que se sucedían de cerca en el oficio, y principalmente en las comunidades de fieles ya llegadas a un grado de perfección, la causa eficiente inmediata de esta institución o determinación del presidente o del mayor, al que llaman *obispo*, igual que la de los menores, a los que llaman *curatos sacerdotes*, y parecidamente, la de todos los otros oficios menores, es o debe ser toda la multitud de aquel lugar con su elección o voluntad expresa, o bien aquél o aquéllos a quien o a quienes la dicha multitud hubiera antes concedido la autoridad de estas instituciones; y que a la misma autoridad pertenece remover o privar de sus oficios a cualquiera de estos dichos oficiales, o compelerles a tomar el oficio dicho si pareciese conveniente.

§ 9. De aquí aún quiero concluir necesariamente que en las comunidades de fieles ya perfectas, pertenece al legislador humano solo, o a la multitud fiel de aquel lugar, sobre la que ha de cuidar el ministro que va a ser promovido, elegir, determinar y presentar personas que se han de promover a los órdenes eclesiásticos, y que a ningún sacerdote u obispo en singular, ni a cualquier colegio de ellos sólo, le es lícito cooperar a la recepción de semejantes órdenes sin la licencia del legislador humano, o del que gobierna con autoridad de él.

[...]

§ 11. Ahora con razón probable, si vale decir probable a lo necesario, quiero mostrar que la elección y aprobación de cualquiera promoviendo a un orden sagrado, pertenece al juez según la tercera significación, a saber, a la sentencia del humano legislador en las comunidades de fieles ya perfectas, y con esto justamente su institución secundaria, con la cual se pone un

⁶ I Cor., 3, 1-2.

obispo o un curato al frente de un cierto pueblo fiel en un cierto lugar, y así en los demás oficios eclesiásticos menores, de la cual sentencia viene también la remoción o privación, y también, si convenga, el constreñir a los ministros eclesiásticos a ejercer su oficio. Luego se mostrará a quién o a quiénes pertenezca el distribuir los bienes temporales eclesiásticos llamados *beneficios*.

Lo primero puede demostrarse por las mismas o parecidas pruebas, con las que en el XII, XIII y XV de la Primera Parte se mostró que la legislación y la institución de los gobernantes pertenecía a la universalidad de los ciudadanos, sólo cambiando la menor de las demostraciones, de modo que la elección, la aprobación de la persona promovida al orden sagrado, y la institución o determinación de la misma para la presidencia de una cierta plebe y provincia, igual que su privación o remoción de la misma, por un delito o por otras razonables causas, se tomen en las demostraciones por el término de la ley o del gobernante.

Y hay en esto una necesidad tanto más evidente de hacerse tales cosas por el legislador o la universalidad de los ciudadanos, cuanto que es más peligroso el errar sobre la persona que se ha de constituir en el grado sacerdotal o en otro alto grado eclesiástico o en el oficio de presidir, más que acerca de la ley humana o de instituir según ella al que ha de tomar las riendas del gobierno. Pues si es promovido al sacerdocio el depravado de costumbres o el ignorante, o el que falla en ambas cosas, y así se le prefiere para cuidar y dirigir al pueblo fiel, de ahí se hace inminente para el pueblo el peligro de la muerte eterna y de muchos inconvenientes civiles; de muerte eterna, porque le toca señalar y dirigir en aquellas cosas y sobre aquellas cosas que tocan a la necesidad de salud eterna.

[...]

§ 14. Ni obsta a esto que los sacerdotes o su colegio, en lo que respecta a la suficiencia de los promovendos, tanto para el sacerdocio como para el oficio pastoral y otros menores, tengan más elementos de juicio; en la misma línea están otras objeciones parecidas aducidas en el XIII de la Primera Parte, con las que parecía argüirse que de ningún modo le compete a la totalidad la legislación o la designación de los gobernantes. Se puede, en efecto, responder a las mismas objeciones de algún modo semejante. Porque demos que los sacerdotes tengan de los tales un más amplio y cuerdo juicio que el que tiene la restante multitud de ciudadanos, lo cual, sin embargo, falla, sobre todo en estos días, las más de las veces, pero no vale concluir de esto que sólo el colegio de sacerdotes tiene de ellos un juicio más cierto que la multitud, de la que forma parte. De donde la restante multitud, junto con el colegio sacerdotal, tendrá un juicio más

cierto y más seguro que el solo colegio de los sacerdotes tomados aparte. Porque *el todo es mayor que una cualquiera parte suya tomada por separado*.

Pero habrá que tener esto bien presente, que la ley bien dada, conforme a la ley divina, debe establecer que el gobernante debe en esto fiarse del juicio de los sacerdotes y de los doctores de la ley divina y de otros varones honestos, lo mismo que en orden a promover en las otras disciplinas, tanto en lo que toca a la enseñanza como en lo que toca a las costumbres, ha de servirse máximamente del juicio de los peritos y hombres probados, digo peritos según la primera significación del juez y de juicio. Porque con el juicio en su tercera significación, bajo la autoridad del legislador, tiene el gobernante la facultad de aprobar o reprobar las personas, establecerlas y removerlas del ejercicio de los oficios, como se demostró en el 15° de la Primera Parte. De lo contrario habría en una misma comunidad tantos gobernantes supremos cuantos son los que juzgan, según la primera significación de juicio, sobre la competencia o incompetencia en cada uno de los oficios de los ciudadanos, lo que habrá que poner entre las cosas inconvenientes e imposibles, si se quiere que permanezca y sea bien regida la ciudad como lo hemos demostrado en el XI de ésta y en el XVII de la Primera Parte.

§ 15. Se han de aprobar, pues, o reprobar con la sentencia o el juicio, según su tercera significación, del legislador, o con la del gobernante con autoridad del legislador, las personas promovendas a las órdenes eclesiásticas, y establecerlas o removerlas de la cura o prelatura mayor o menor y prohibirles su ejercicio, y también, si por malicia se abandonasen en dicho ejercicio, obligarlas a ejercerlo, no sea que por su perversidad alguien corriera el peligro de la muerte eterna, como por falta del bautismo o de otro sacramento. Lo cual se ha de entender de las comunidades ya perfectas. Porque en aquel lugar donde el legislador o el gobernante investido de su autoridad fueren infieles, como ocurría en muchas comunidades y en casi todas en el estado de la iglesia primitiva, el aprobar y el reprobar las personas promovendas a los órdenes eclesiásticos, con todas las instituciones ya dichas y ejercicios de oficios, esta autoridad recaería en el sacerdote o el obispo con la parte más sana de la multitud de fieles existentes allí, o del mismo, sólo si está solo allí, sin el consenso o conocimiento del gobernante, para que por aquella promoción e institución de prelados o curatos la fe y la doctrina saludable de Cristo se propagaran, lo que no se haría con la autoridad, empeño y precepto de un legislador infiel o custodio de la ley, más bien se impediría. De este modo hicieron los apóstoles a los comienzos de la iglesia de Cristo y eran obligados a hacerlo por precepto divino y

estarían obligados sus sucesores a lo mismo a falta del legislador. De donde el apóstol, en la 1ª a los Corintios, 9º: *Pues si evangelizo no es mía la gloria, pues me urge una necesidad; ¡ay de mí si no evangelizaré!*⁷ Pero donde el legislador fiel y el custodio de la ley quieren que haga eso, digo que de ellos es la autoridad según el modo dicho, en virtud de las razones y probaciones dadas, sacadas ya de la Sagrada Escritura, ya de la razón humana, probable y necesaria.

§ 16. En lo tocante a la distribución de los bienes temporales que suelen denominarse *beneficios eclesiásticos*, conviene saber de antemano que tales bienes temporales o están destinados por el legislador para la sustentación de los ministros del evangelio y de todas aquellas personas necesitadas, de las que dijimos en el XIV y XV de esta Parte, o son ordenados para este uso por alguna persona o algún colegio particular. Si tales bienes temporales fueron así destinados provenientes de la donación y la autoridad del legislador, digo que puede lícitamente confiar la autoridad de su distribución, según la ley divina, a quien y cuando le plazca, y por alguna causa, retirar esa autoridad de aquél o aquéllos a los hubiere confiado tal autoridad; y lo opuesto no puede ser probado por la Escritura, sino más bien lo propuesto; como también por la autoridad de Ambrosio, en la carta *De la cesión de las basílicas*, lo mostramos antes, en el XIV de esta Parte. No sólo puede el legislador fiel, lícitamente según la ley divina, revocar la autoridad de distribuir tales bienes temporales de aquél o aquéllos a los que lo había confiado, sino también venderlos, o de otro modo enajenarlos por alguna causa razonable ocurrente, pues suyos son y están siempre en su poder y derecho, a no ser que acaso la misma cosa simplemente, o también con su dominio, la hubiere transferido a la potestad de algún colegio o persona particular, añadiendo, no obstante, en todo casos que a todo evento o circunstancia el pueblo fiel está obligado por ley divina, si tiene para ello, a la sustentación en comida y vestido decente de los ministros del evangelio, con los que deben éstos contentarse, como de la 1ª a Timoteo, último⁸, se demostró en el XIV de esta Parte. Mas si los bienes temporales se hubieren asignado a pías obras, por donación o por legado de una persona o de personas particulares, digo que los tales bienes deben conservarse, guardarse y distribuirse según la intención del donante o legante, por el humano legislador o por gobernante con su autoridad.

[...]

§ 18. También de lo dicho antes no puede ocultársenos que el humano legislador, o el que tiene el gobierno con su autoridad, puede lícitamen-

⁷ I Cor., 9, 16.

⁸ I Tim., 6, 8

te, según ley humana y divina, imponer tasas y censos sobre los bienes eclesiásticos, sobre todo de los provenientes de los bienes inmuebles, si sobrare algo de lo necesario para la suficiencia de los ministros evangélicos, bienes que dedicará a la defensa de la patria y a la redención de los cautivos en obsequio de la fe, o a soportar las cargas públicas y otras razonables causas, según la determinación del legislador fiel. Porque el que tales bienes temporales estableció por donación o legado para causas pías y confió para su distribución a alguien o algunos, no pudo entregarlos a un colegio o a una persona particular con mayor inmunidad de tributos que la que tenían antes cuando estaban en su potestad. Pero para aquel tiempo nunca fueron inmunes de las cargas públicas, luego ni después que por el donante o el instituyente fueron entregados a la potestad de algún otro.

[...]

XVIII. DEL ORIGEN Y ESTADO PRIMERO DE LA IGLESIA CRISTIANA Y DE DONDE
EL OBISPO Y LA IGLESIA ROMANA TOMARON LA AUTORIDAD PREDICHA
Y UN CIERTO PRIMADO SOBRE LOS DEMÁS

Ahora nos resta, a partir de las intenciones propuestas, poner de manifiesto el comienzo y origen de donde provino a algunos obispos y presbíteros la jurisdicción coactiva y el de todas las instituciones sacerdotales denominadas no esenciales, y la potestad de distribuir todos los bienes temporales eclesiásticos, y de donde se adscriba a sí el Papa romano la suprema de estas tales potestades. Y a continuación de esto, añadir a quién o a quiénes les competa la justa potestad de interpretar los sentidos dudosos de la Sagrada Escritura y de proponerlos e imponerlos así interpretados a los fieles para ser creídos y observados.

[...]

§ 3. Poniendo, pues, el comienzo en el sagrado canon, como la fuente de la verdad buscada, tomemos la palabra de Cristo, en Juan, 20º, por la que entrega a los apóstoles la autoridad sacerdotal o la potestad de las llaves, una y otra indistintamente cuando, exhalando sobre ellos el aliento, les dijo: *Recibid el Espíritu Santo. A los que perdonareis los pecados les serán perdonados*¹; añadiendo a esto el mandato que les impuso también de predicar indiferentemente el evangelio por todo el orbe, cuando les dijo, en Mateo, último: *Id y haceos discípulos de todas naciones*, etc.² A ellos, por inmediata vocación suya, les añadió después a Pablo, como vaso de elección, a saber, del Espíritu Santo, como de los Hechos, 9º, aparece³.

¹ Juan, 20, 22.

² Mt., 28, 19.

³ Hechos, 9, 15.

Este y juntamente los otros apóstoles, poniendo en ejecución el dicho mandato, como se ve claro por los Hechos y por las epístolas, predicaron y enseñaron el evangelio o la fe cristiana, a la cual se adhirieron muchos allí; luego por revelación divina y por ordenación de ellos entre sí, unos se quedaron en Judea, otros se trasladaron, separándose, a diversos pueblos y provincias. En las cuales predicando con libertad y constancia el evangelio, progresó, con la conversión de personas de uno y otro sexo, cada uno, cuanto pudo y le dio el Señor, como se lee en sus relatos y en sus historias comprobadas si bien, entre todos, con su predicación, progresaron notablemente dos, a saber, los bienaventurados Pedro y Pablo. [...].

§ 4. De estos dos apóstoles Pedro y Pablo, por su mayor parte, aunque según el pasaje del canon, de Pablo más principalmente por lo que toca a los gentiles, se derivó el ritual de la iglesia cristiana; porque el mismo Pablo fue el principal y el primer apóstol enviado a las gentes, o a los gentiles, o al prepucio, como Pedro a la circuncisión, como aparece por el 2º a los Gálatas y a los Romanos, 11º, y también por la glosa de los santos en aquel pasaje y en otros muchos lugares de la Sagrada Escritura.

§ 5. Así pues, estos dos apóstoles con todos los demás, imitando a su maestro, a Cristo, vivieron bajo la jurisdicción coactiva de sus príncipes de este siglo y enseñaron a otros a vivir así, como abiertamente se ha demostrado en el IV y V de esta Parte. Y así vivieron sus sucesores, presbíteros u obispos, con sus diáconos y otros ministros del evangelio, sucesivamente, hasta los tiempos del primer Constantino, emperador de romanos, como del susodicho Código de Isidoro bastante claro aparece, de manera que ninguno de los obispos por todo aquel tiempo ejerció sobre los otros obispos una jurisdicción coactiva; si bien muchos obispos de otras provincias, en las cosas que dudaban, ya de la Escritura Sagrada, ya del ritual eclesiástico, no atreviéndose a congregarse públicamente, pidieron consejo al obispo y a la iglesia de fieles existente en Roma, quizá por la mayor multitud allí de fieles y más expertos, por el hecho de que entonces estaban en auge allí los estudios de todas las ciencias, de donde sus obispos y sacerdotes eran más impuestos en el saber, y de los tales había en su iglesia más abundancia que en las otras. Eran también mirados con más respeto, ya porque el bienaventurado Pedro era el de más edad de los apóstoles, estaba más adornado de méritos y era más respetable; allí se dice que tuvo su sede episcopal, igual que el bienaventurado Pablo, del que consta más ampliamente, como quedó claro por el XVI de esta Parte; ya también en razón de la principalidad de la urbe y su celebridad comparada con las otras provincias del mundo. De donde también los fieles de otras provincias, careciendo de personas competentes para gobernar sus iglesias, pedían al obispo y a la iglesia de fieles

de Roma personas para ponerlas al frente del obispado, por el hecho de que la iglesia de Roma ofrecía de tales personas, como ya dijimos, mayor abundancia. Y los obispos y la iglesia de los romanos, a quienes se demandaba consejo y auxilio, tanto acerca de la fe, como de lo tocante al ritual eclesiástico y provisión de personas, caritativa y fraternalmente atendían en esto a los indigentes y suplicantes, a saber, enviando obispos cuando apenas se hallaban quienes aceptaran serlo. Y sus ordenaciones, las que dictaban sobre el ritual eclesiástico, amigablemente las comunicaban con otras provincias, y algunas veces las atendían amonestando con amor cuando oían que en otras provincias había surgido alguna contienda o cisma.

[...]

§ 7. De esta así dicha cuasi consuetudinaria prioridad, con consenso espontáneo de las otras iglesias, los obispos de los romanos, prolongando el proceso del principio, fueron tomando una cierta autoridad, cada vez mayor, de constituir decretos u ordenaciones sobre la universal iglesia en lo tocante al ritual eclesiástico y en los actos también de los sacerdotes, y preceptuando la obligación de su cumplimiento; esto hasta los tiempos del dicho Constantino. Pero la cuestión de si esto lo pudieron hacer los obispos de los romanos por su sola autoridad, o si fue preciso que se añadiera a ello el consentimiento de los otros, en lo que sigue iremos diciéndolo. Pues el predicho Constantino, como lo narra Isidoro en el ya dicho Códice, en el capítulo De la iglesia primitiva en el Concilio Niceno⁴, y se contiene en el edicto del señor Constantino emperador, fue el primer emperador que, por el ministerio del bienaventurado Silvestre, entonces Papa romano, públicamente abrazó la fe, y el que parece que, el primero, eximió al sacerdocio de la jurisdicción coactiva de los príncipes, el que también, por el mismo edicto dicho, parece ser quien primero atribuyó a la iglesia de Roma y a su obispo autoridades y poderes sobre todos los otros obispos e iglesias, que ahora afirman serles debidos por otra razón, como dejamos expuesto en el XIX de la Primera Parte, párrafos 8 y 9; y con esto, las jurisdicciones coactivas, sobre las mismas tierras, fincas y posesiones numerosas, con un cierto dominio secular incluso sobre algunas provincias, como claramente advierte el que pasa la vista por dicho edicto. Este fue también el primer emperador, como se lee arriba, que dio permiso a los cristianos para congregarse públicamente, para levantar templos o iglesias, y con cuyo precepto tuvo lugar el primer concilio congregado en Nicea.

[...].

§ 8. Así pues, una vez expuesto el proceso de las cosas que vamos a aclarar, tomándolo desde su origen, afrontando con más detalle su determi-

⁴ Seudo-Isidoro, *Decretales* (ed. Hinschius), 248.

nación, pondremos como base, con el apóstol, la suposición libre de toda duda de que la fe católica es una, no muchas; de donde a los Efesios, 4º: *Un Señor, una fe*⁵. Y que en el mismo sentido, en la unidad, ha de ser creída y confesada por todos los fieles, como allí, poco más abajo, dice el apóstol: *Hasta que lleguemos todos a la unidad de la fe y al conocimiento del Hijo de Dios*⁶. De lo cual inferimos necesariamente, primero, que si hubieren surgido sentidos y sentencias dudosas de la ley divina, máxime la evangélica, y si entre los doctos se hubieran levantado contiendas y controversias probables, como por la ignorancia o la malignidad de algunos, o por las dos cosas, leemos que surgieron ya, según la profecía de Cristo y del apóstol, era preciso zanjarlas. Pero el conseguirlo mostraremos que pertenece necesariamente sólo al concilio general de todos los fieles, o de aquellos que tuvieren la autoridad de todos los fieles; a ellos sólo compete esa determinación.

Después mostraré, según la ley divina y según la recta razón, que el convocar el concilio general y, si es preciso, congregarlo con poder coactivo, pertenece a la autoridad del legislador fiel no sometido a otra autoridad mayor, no a alguna persona singular o a algún colegio particular, cualquiera que sea su dignidad o condición, a no ser que al mismo o a los mismos les haya sido otorgada la autoridad por el legislador general supradicho.

Además mostraré con certeza que nada se puede establecer acerca del ritual eclesiástico y de los actos humanos que obligue a todos los hombres a su observancia bajo alguna pena para el estado presente o futuro, a no ser por el solo concilio general, o inmediatamente por el supremo legislador fiel, o por autoridad derivada de él. Por lo que también mostraremos a continuación que ningún príncipe, provincia o comunidad puede ni debe ser sometida a entredicho o a excomunión por un sacerdote u obispo, quienquiera que sea, si no es según el modo ordenado por la ley divina, o por el supradicho concilio general.

Después se mostrará evidentemente que no pertenece a la autoridad de ningún único obispo ni de otra persona singular o colegio particular constituir a ningunas personas en todos los oficios eclesiásticos del mundo, ni distribuir o conferir a ellos los bienes temporales eclesiásticos llamados beneficios; sino que la autoridad ésta es sólo del fundador o del donante o del legislador universal fiel, o de aquél al cual o a los cuales, y según la forma y modo, en que el mismo fundador o legislador hubiere otorgado esa potestad.

⁵ *Efes.* 4, 5.

⁶ *Ibid.*, 4, 16.

Después se mostrará que es conveniente instituir un único obispo y una iglesia, los cuales, por su autoridad, sean, ella más principal y como cabeza de las demás, y a él pertenezca, con su iglesia, sugerir a todos los otros obispos e iglesias aquellas cosas que hubieren sido ordenadas por los concilios generales tocante al ritual eclesiástico y a otros actos humanos para la común utilidad de los fieles, y que parezcan dignas de ser ordenadas.

Finalmente, de todo esto inferiremos necesariamente que, tanto lo determinado acerca de la Sagrada Escritura y fe católica, como lo referente al ritual eclesiástico, con todas las demás cosas determinadas por el concilio general, sólo por la autoridad del concilio general, no por la de otro cualquier colegio o persona singular alguna, podrán ser inmutadas, disminuidas o suspendidas, o totalmente revocadas.

De las cuales cosas todas vendrá al conocimiento sensato de cada cual, que el obispo romano, o su iglesia, o cualquier otro obispo o iglesia suya, en cuanto tales, ninguna potestad o autoridad de las ya dichas ejercen sobre los demás obispos e iglesias por derecho divino o humano, sino solamente aquella que absolutamente o para algún tiempo le hubiere sido concedida por el supradicho concilio general. Por lo cual también queda manifiesto que el obispo romano o cualquiera otro impropriamente, menos debidamente y fuera, más aún, contra el sentir de las escrituras divinas y de las ordenaciones humanas, se atribuye a sí la plenitud de potestad sobre el príncipe, o sobre una comunidad, o sobre una persona singular; y que el tal obispo y otro cualquiera ha de ser totalmente impedido en esa atribución, aun por amonestación y coactiva potestad, si es preciso, ejercida por los legisladores humanos o por los que bajo su autoridad gobiernan.

[...]

XX. A QUIEN PERTENEZCA O PERTENECIÓ LA AUTORIDAD DE DEFINIR
O DETERMINAR LOS SENTIDOS DUDOSOS DE LA SAGRADA ESCRITURA

Adelantadas así estas premisas, resumiendo lo que nos propusimos concluir, queremos mostrar, lo primero, que a los sentidos o enunciados dudosos de la Sagrada Escritura surgidos, o por surgir cuando surjan, principalmente en lo tocante a los artículos de la fe, preceptos o prohibiciones, es conveniente y necesario ponerles un término. Porque es conveniente, más aún, necesario, aquello sin lo cual no se salvaría de ningún modo la unidad de la fe, sobrevendría error o cisma entre los fieles de Cristo.

Y esto es la determinación en las opiniones dudosas y a veces contrarias de algunos doctores acerca de la ley divina. Porque las diversas

opiniones y posiciones contrapuestas en torno a ésta, induciría diversas sectas, cismas y errores, como se narra en el ya mencionado Códice de Isidoro, capítulo titulado: *Comienza el Prefacio al Concilio de Nicea*¹.

[...]

§ 2. A seguido de esto muestro que la autoridad principal de esta determinación, mediata o inmediata, es exclusivamente el concilio general de los cristianos, o de la parte prevalente de los mismos, o de aquellos a los que esta autoridad les hubiere sido concedida por la universalidad de los fieles cristianos; de tal manera, en efecto, que todas las provincias del mundo o comunidades notables, según lo determine el legislador humano, ya sea éste uno único o muchos, y según la proporción de las mismas en cantidad y en cualidad de las personas, elijan varones fieles, primero presbíteros, y luego también no presbíteros, pero idóneos como más probados en su vida y más expertos en la ley divina, los cuales como jueces, según la primera significación de juez, representando las veces de toda la universalidad de los fieles, con la autoridad ya dicha concedida a ellos por las universalidades particulares, se congreguen en un determinado lugar del orbe, el más conveniente según la decisión de la mayor parte de, ellos, en el cual definan conjuntamente aquellas cosas tocantes a la ley divina que se hubieran presentado como dudosas y como con utilidad, conveniencia y necesidad de ser de terminadas, y con ello también lo que sobre lo demás del ritual eclesiástico o culto divino, o lo que para el futuro, para la paz y tranquilidad de los fieles, se ha de ordenar. Porque sería por demás inútil que se reuniera para la tal congregación la multitud de fieles inexperta; inútil también porque se vería turbada por las obras necesarias para la sustentación de la vida corporal, lo que sería para ella oneroso y acaso intolerable.

[...]

§ 4. Y que competa sólo al concilio general ya dicho la autoridad predicha de definir y ordenar según el modo apuntado, y a ninguna otra persona singular o colegio, se puede demostrar con autoridades de la Sagrada Escritura y con razones parecidas a aquéllas con las que probamos en el XII de la Primera Parte y en el XVII de esta Segunda, que pertenecía a él la legislación y la institución secundaria de los oficios eclesiásticos, sólo mudando el término menor de las demostraciones, de modo que las cosas dudosas tocantes a la ley divina, que hay que determinar y definir, con todo lo demás que hay que regular referente al ritual eclesiástico o culto divino, y para la paz y unidad de los fieles, se pongan, en vez del término ley o institución secundaria de los oficios eclesiásticos. Tanto más necesario es

¹ Seudo-Isidoro, *PL*, 130, 253-254.

poner atención en todo esto cuanto más diligente es la discreción y el cuidado que hay que tener, en lo que, toca a la ley o fe que hay que profesar y a aquellas cosas que pueden aprovechar o dañar mucho a todos los fieles.

§ 5. Porque así lo hicieron los apóstoles con los ancianos sobre aquellas cosas que se ofrecieron dudosas acerca del evangelio, como aparece en los Hechos, 15°, y más de propósito lo adujimos en el XVI de esta Parte. Porque aquella duda sobre la circuncisión no la resolvió el bienaventurado Pedro o algún otro apóstol por sí en singular, sino que se congregaron sobre ello todos los apóstoles y ancianos o los más expertos en la ley. Y señal de que es verdad lo que decimos, es que en los concilios primordiales, congregados para definir los puntos dudosos de la Escritura, estaban presentes los emperadores y emperatrices fieles con sus oficiales, como aparece bastante claro por el mencionado Códice de Isidoro en el pasaje que aducimos en el capítulo siguiente, del párrafo 2 al 8, bien que en aquel tiempo no apretaba tanto la necesidad de la presencia de los no sacerdotes como en el tiempo moderno, por la mayor turba, más de lo conveniente, de sacerdotes y obispos ignorantes de la ley divina. De donde, discrepando entre sí los sacerdotes sobre lo que hay que creer para la salud eterna, es la parte prevalente de los fieles la que ha de juzgar cuál es la parte más sana entre ellos; bien que, puestos de acuerdo en aquello que había de dudoso, habrá que creer, en las cosas dichas, a aquellos que fueron promovidos a las órdenes del modo dicho en el XVII de esta Parte.

§ 6. Por lo que también quiero de nuevo demostrar que tal determinación de ningún modo pertenece a solo el obispo romano, ni tampoco a él junto con su colegio de cardenales y, consiguientemente, tampoco a ningún otro obispo, ni solo ni con algún otro colegio particular. Pongamos el caso, como puede ser que ocurra, que algún hereje es asumido para el pontificado romano, o que después de asumido, si no antes, cae en este mal, por ignorancia o por malicia o por las dos cosas, como de hecho leemos que han sido asumidos algunos, tal Liberio de nacimiento romano². Si, pues, sólo el pontífice o él con sólo su grupo de cardenales, que podrían verosímilmente ser cómplices de su error, porque los toma como los quiere, sin intervención determinante de nadie más, y porque él dice que puede tomarlos, decidiese sobre alguna opinión dudosa de la Escritura, ¿habría que atenerse al sentir de este obispo solo o de sólo su colegio, o de la mayor parte, los que acaso fueron seducidos por ignorancia o por malicia, por codicia o ambición, o por otra cualquiera pasión siniestra?

[...]

² Martinus Polonus, *Chronicon Pontificum et Imperatorum*, Mon. Germaniae Historica. Scriptores, 22, 416, 30.

§ 8. Más, si de sólo el Papa romano, o de otro cualquier obispo solo, fuese tal autoridad, o si, según Isidoro en el predicho Códice, en el capítulo que lleva por título: Prefacio de Isidoro a la obra siguiente³, las cartas o decretos del pontífice romano fueran en autoridad iguales o no inferiores a aquellos que fueron determinados o definidos por el concilio general, todos los principados del siglo, todos los reinos y provincias del mundo y personas singulares de cualquier dignidad, preeminencia o condición que fueren, estarían sometidos con jurisdicción coactiva al que es cabeza en el episcopado romano. Porque Bonifacio VIII, Papa romano, sancionó esto por una cierta carta o decreto, cuyo comienzo es: *Una santa iglesia católica*, y su fin: *Declaramos, pues, pronunciamos y definimos que de necesidad de salud toda humana creatura está sometida al romano pontífice*⁴. Si, pues, lo una vez determinado tocante a la Escritura por el concilio general debidamente convocado, congregado y en la forma debida celebrado y consumado, principalmente aquellas cosas que es necesario creer como verdaderas para la salud eterna, quedan fijadas como una verdad inmutable e infalible, como se demostró al principio del 19º de esta Parte, la carta esta de Bonifacio obtendría una cierta, indudable e irrevocable verdad. Ahora bien, consta que ella, desde el principio, ahora y siempre, es falsa, errónea, y, la más perjudicial, de todas las falsas imaginables, para los que viven en el orden civil, como sin lugar a dudas hemos demostrado en el IV, V y IX de esta Parte.

§ 9. Y que sea falsa, igual que el dicho de Isidoro, a no ser que sea aliviado con una interpretación piadosa, lo pone de manifiesto la carta o decreto de Clemente V, sucesor del mismo Bonifacio, cuyo inicio es: Es mérito... del carísimo hijo nuestro Felipe, rey ilustre de Francia, y su final: Tanto para la iglesia, como para el rey y reino nombrados⁵. Por ésta, en efecto, se expresa el dicho Clemente definiendo que la carta bonifaciana en nada perjudica al rey o reino predicho. No creyendo, pues, ni habiendo de creer el dicho rey, ni el sucesor suyo, ni ninguno de sus súbditos, de acuerdo con la persuasión de Clemente y de su sentir y el de todos los cristianos que es verdadero, sino abiertamente falso, aquello que contiene, la carta supradicha en aquel artículo en el que somete así todos los principados y reinos, síguese necesariamente que creer eso no es de necesidad de salud; si lo fuera, sería perjudicial al que no lo creyera.

³ Seudo-Isidoro, *Decretales* (ed. Hinschius), p. 8.

⁴ Bula: *Unam sanctam*. *Corpus Iuris Canonici, Extravag. Commun.* I, tít. VIII, c. 1 (18 nov. 1302).

⁵ Bula: *Meruit*. *Corpus Iuris Canonici, Extravag. Commun.*, l. V, tít. VIII, c. 2 (febrero, 1306).

XXI. A QUIEN PERTENEZCA LA AUTORIDAD COACTIVA DE CONGREGAR EL CONCILIO GENERAL DE SACERDOTES Y OBISPOS Y OTROS FIELES, Y A QUIEN PERTENEZCA LA AUTORIDAD, EN EL MISMO CONCILIO, DE ESTABLECER LO QUE OBLIGUE A LOS FIELES BAJO PENA O CULPA EN EL SIGLO PRESENTE O EN EL FUTURO, Y A QUIEN DE NUEVO PERTENEZCA CASTIGAR EN EL MUNDO PRESENTE A CUALQUIER TRASGRESOR DE LO ESTABLECIDO O DEFINIDO EN EL CONCILIO GENERAL. TAMBIÉN QUE NINGÚN OBISPO O SACERDOTE PUEDE EXCOMULGAR AL PRÍNCIPE NI PONER ENTREDICHO A NINGÚN PUEBLO, NI CONFERIR A NINGUNO BENEFICIOS TEMPORALES ECLESIASTICOS, NI DIEZMOS, NI LICENCIAS DE ENSEÑAR, NI OFICIOS CIVILES DE CUALQUIER GÉNERO, SINO SEGÚN LA DETERMINACIÓN Y CONCESIÓN DEL CONCILIO GENERAL O DEL LEGISLADOR HUMANO, O DE AMBOS

[...]

§ 13 Además. Concedida esa potestad general e inmoderada sobre la institución de oficios y la distribución de los bienes temporales, o beneficios, todos los reinos y comunidades políticas quedan expuestos al peligro de la disolución, o de una inmensa perturbación, si el tal obispo pretende someter a sí a los principados seculares, los cuales todos, según la invención escrita llamada *decretal*, que ya adujimos más arriba, proclamó Bonifacio VIII, están sometidos a cualquier obispo romano con jurisdicción coactiva, de tal manera que aseveró pertinazmente que creerlo así los fieles era *de necesidad de salud*. Y lo mismo puede estar implicado en los edictos de un así dicho obispo romano contra el renombrado Luis, de los duques de Baviera, elegido rey, aunque esto sólo parezca referirse al reino o imperio de los romanos; pero comprendiendo a todos los demás en virtud de aquel título que, mezclado con otras razones, se atribuye, a saber, *la plenitud de potestad*; porque con este título no se hace meramente superior, con jurisdicción coactiva, al príncipe de los romanos, sino también a cualesquiera otros reyes, como con evidencia lo mostramos en el capítulo último de la Primera Parte.

XXII. EN QUÉ MODO EL OBISPO ROMANO Y SU IGLESIA SEAN CABEZA Y LA PRINCIPAL ENTRE LAS DEMÁS; Y EN QUÉ CLASE DE AUTORIDAD LES CONVENGA ESTO

[...]

§ 18. Así pues, el obispo romano y su iglesia, desde el principio, comenzaron a obtener lícitamente esa primacía por cierta solicitud suya caritativa que después, por la devoción, reverencia y obediencia, convertidas en costumbre, adquirió la fuerza de una elección. Porque nunca de la

Escritura se podrá colegir que por precepto o consejo de Cristo o de algún apóstol las otras iglesias u obispos hayan de someterse a la iglesia o al obispo de los romanos, aun en el ritual eclesiástico. Pero si hubiese sido debido de necesidad de salud para los fieles, como afirman ya algunos obispos romanos, ni sólo en el orden del ritual eclesiástico, sino también con jurisdicciones coactivas, y no sólo para los clérigos, sino hasta para cualesquiera principados seculares, ¿cómo se explica que Cristo y sus apóstoles hubieran pasado por alto el transmitirlo? Pero como tanto Cristo como los apóstoles abiertamente estatuyeron lo opuesto, sobre todo en lo tocante a las jurisdicciones coactivas, como de la Escritura se ha demostrado claramente en el IV, en el V y en el IX de esta Parte, las palabras de los que tal dicen han de computarse como fábulas apócrifas.

§ 19. Pero en el tiempo de Constantino emperador romano, que recibió públicamente la fe de Cristo y, el bautismo, los fieles comenzaron a reunirse a plena luz, a definir lo dudoso tocante a la fe y a ordenar el ritual eclesiástico, como aparece del citado Códice de Isidoro, en el capítulo: *De la primitiva iglesia, en el sínodo niceno*¹.

De este Constantino, pues, por edicto imperial, acomodándose a la antedicha laudable y antigua costumbre, obtuvo el obispo y la iglesia de los romanos la primacía que hemos precisado convenirle sobre las otras, y a más de esa primacía, posesiones y dominios de algunas provincias, bien que antes de los tiempos de Constantino, y aun después, algunos obispos romanos, en sus cartas y en algunos decretos sugirieron que la primacía sobre los demás que, según hemos demostrado, les vino por elección o por instituciones de los príncipes, se les debía a ellos en singular por ley divina, sin pedirla ellos ni serles concedida por el legislador humano fiel, ni por algún otro colegio o persona particular de cualquier preeminencia o autoridad que fuera; lo opuesto de lo cual hemos demostrado suficientemente en el capítulo precedente.

§ 20. Después de los tiempos de Constantino I, y, sobre todo, en la sede vacante del imperio, algunos obispos de romanos, en sus cartas, dieron a entender que aquella primacía se les debía, unas veces por la ley divina, otras por concesión de los príncipes. Mirando a la extensión y cualidad con que se había de tomar dicha primacía, muchos de ellos sugerían y en algún modo expresaban que abarcaba la interpretación de la ley divina y la ordenación del ritual eclesiástico, tanto sobre el culto divino, como sobre los ministros, todo lo que se refiere a su institución inseparable y primaria que denominamos esencial, y la secundaria, y a la privación de la misma, la que

¹ Seudo-Isidoro, *Decretales*, ed. c., p. 247.

llamamos separable o accidental en el XV de esta Parte. Extendieron también ellos esa (autoridad) a todos los obispos y a todas las iglesias, pueblos y personas particulares, a la sentencia de excomunión y de entredicho de los divinos oficios, tanto para súbditos como para ministros o prelados de las iglesias, y a otra cualquiera censura eclesiástica semejante, como el anatema, que se hace recaer sobre los dichos fieles. Otros después pretendieron lo mismo tocante a toda jurisdicción o potestad coactiva sobre todos los ministros y colegios de ellos de todas las iglesias del mundo. Y usaron de ella como perteneciendo ellos por concesión de los príncipes, mientras se mantuvo poderoso el príncipe de los romanos y su reino íntegro y llena su sede. Pero venido este reino a su disolución y, sobre todo, durante la vacante de la sede imperial, a tiempos, se hicieron fuertes en esta autoridad como concedida por ley divina. Cómo y de qué diverso modo ocurrió esto lo diremos en el capítulo siguiente.

Así también declararon en sus cartas que les pertenecía el reparto y la distribución de todos los bienes temporales eclesiásticos conforme a su deseo, sin demandarlo ni pedir el consenso de ningún colegio o persona particular, de cualquier dignidad y autoridad que fuese. Y no contentos con estos excesos, algunos de los más modernos afirmaron en sus cartas o decretales que se les debía por ley divina la autoridad o la jurisdicción coactiva suprema sobre todos los principados del mundo, pueblos y personas particulares, de modo que ninguno de los gobernantes ya dichos pueda, sin, o contra, su consentimiento o dictamen, ejercer lícitamente la jurisdicción coactiva, que llaman ellos la *espada temporal*, y a los que obran al margen, o en contra, de esto, gobernantes y pueblos, los declaran sujetos a la sentencia de excomunión o de entredicho pronunciada de viva voz. Se autodefinen a sí solos vicarios de Cristo en el mundo, de Cristo que *es rey de reyes y señor de los que dominan*², arrogándose con ello solapadamente el título, como debido a ellos, de la *plenitud de potestad*. Por lo cual dicen también pertenecer a su autoridad el poder de dar y quitar lícitamente los reinos todos y principados del mundo a los reyes y a los otros gobernantes, quitarlos a los trasgresores de sus mandatos, aunque sean éstos muchas veces, con arreglo a la verdad, impíos e ilícitos.

Entre los otros obispos romanos, Bonifacio VIII, no menos temeraria que perjudicialmente, y contra el sentido literal de la Escritura, apoyado en interpretaciones fantasiosas de ella, a tanto llegó en sus afirmaciones y declaraciones, que decretó que es de necesidad para la salud, eterna que se crea y reconozca por todos que tal potestad es debida a los obispos romanos. Opinión que siguieron su sucesor Clemente V y el que se hace pasar

² Apoc., 19, 16.

por sucesor de este Clemente, bien que parece se refieren explícitamente sólo al imperio de los romanos. Y haciéndolo como lo hacen, apoyados en el supradicho título, a saber, de la plenitud de potestad dada a ellos por Cristo, no cabe duda de que dicha potestad o autoridad, si alguna se deduce de aquello, abarca igualmente a todos los reinos y principados del mundo, como la potestad de Cristo, como más de propósito lo mostramos en el 19º y último de la Primera Parte.

XXIII. DE LOS MODOS DE LA PLENITUD DE POTESTAD, Y CON ARREGLO
A QUÉ MODO Y ORDEN SE LOS HA ATRIBUIDO A SI EL OBISPO ROMANO
Y SUMARIAMENTE COMO HA USADO Y USA DE ELLOS

[...]

§ 3. Mas como la plenitud de potestad parece importar una cierta universalidad, y nuestro propósito es tratar sólo de las potestades voluntarias, hemos de distinguir la plenitud de potestad en sus modos, fijándonos en la diferencia de potestad universal voluntaria.

Es, y se puede con verdad entender de un modo la plenitud de potestad, según la significación y fuerza del lenguaje, aquella que, sin tope ninguno, es capaz de cualquier acto posible y de realizar voluntariamente cualquier cosa. La cual, entre los hombres parece convenir sólo a Cristo, de donde en Mateo, último: *Se me ha dado toda potestad en el cielo y en la tierra*¹.

De un segundo modo, y más a nuestro propósito, puede entenderse plenitud de potestad aquélla según la cual es lícito al hombre ejercer cualquier acto suyo voluntario imperado sobre cualquier hombre y sobre cualquier cosa externa que esté en poder de los hombres, o sea ordenable a su utilidad, o aquélla según la cual haya un poder total sobre dicho acto, aunque no sobre cualquier hombre o cualquier cosa sometida a los hombres, o también aquélla según la cual no le sea lícito ejercerse sobre cualquier acto, sino sobre alguno determinado en su especie y modo, pero según todo el ímpetu de la voluntad del sujeto referida a cualquier hombre o a cualquier cosa sometida a la humana potestad.

De un tercer modo se puede entender plenitud de potestad aquella que tiene la jurisdicción coactiva sobre todos los principados, pueblos, comunidades, colegios y personas particulares, o también sólo sobre algunos de ellos, pero con todo el ímpetu de la voluntad.

¹ Mt., 28, 18.

De un cuarto modo puede entenderse la ya dicha, o según el modo dicho, pero ejercida sólo sobre todos los clérigos, y la de instituirlos para los oficios eclesiásticos, privarlos o deponerlos de ellos, y de distribuir los bienes temporales o beneficios, o según el modo antes dicho.

De un quinto modo puede entenderse aquella que es de los sacerdotes, de atar y desatar completamente a los hombres de sus culpas y penas, de excomulgar, de poner en entredicho y de reconciliar, de la que dijimos en el VI y VII de esta Parte.

De un sexto modo puede entenderse aquélla por la que puede imponerse las manos a los que se han de tomar para los órdenes eclesiásticos, la de administrar o denegar los sacramentos, de la que hablamos en el XVI y XVII de esta Parte.

De un séptimo modo puede entenderse aquella que es de interpretar los sentidos dudosos de la Escritura, principalmente en aquello que es de necesidad de salud, definir y distinguir lo verdadero de lo falso, lo cuerdo de lo disparatado y de ordenar todo el ritual eclesiástico y emitir precepto coactivo, o generalmente bajo anatema, precepto de la observancia de lo ordenado.

De un modo octavo y último, por lo que hace a nuestro propósito, puede entenderse la plenitud de potestad aquella que es de la cura pastoral de las almas, en general, abarcando todos los pueblos y provincias del mundo, de la que dijimos en el IX y el XXII de esta Parte.

Podría también entenderse la plenitud de potestad, cada una de las que hemos distinguido, que no está determinada por una ley, y no plena la que está determinada por una ley humana o divina, bajo la cual puede también convenientemente comprenderse la recta razón.

§ 4. Existen posiblemente otros modos y combinaciones de la plenitud de potestad, pero creemos que hemos enumerado todos los que hacen a nuestro propósito e intención.

Adelantadas, pues, estas divisiones de plenitud de potestad, digo que la plenitud de potestad según los dos primeros modos, no conviene al obispo romano ni a ningún otro sacerdote o no sacerdote fuera de Cristo o Dios. Los paso por alto por su evidencia y por estar testificados por la sabiduría divina y humana y toda ciencia moral, y por razón de brevedad.

Sobre el tercer y cuarto modo de la plenitud de potestad se probó por demostración en el XV de esta Parte, por testimonio infalible de la Sagrada Escritura en el IV, V y VIII de esta Parte, todavía confirmado en el XV, XVI y XVII, y finalmente corroborado en el XXI de la misma, que a ningún sacerdote u obispo, en cuanto tales, les compete por ley divina en absoluto, y menos con alguna plenitud, ejercerla sobre ningún clérigo (o no

clérigo). Pero si por humana ley le ha sido concedida tal plenitud de potestad a algún clérigo, obispo, sacerdote o no sacerdote, según algún modo posible de concederla y de revocarla, con causa razonable, por el juicio del legislador humano, hay que verificarlo a partir de las humanas leyes, rescriptos o privilegios del mismo legislador.

De los otros modos, el quinto y el sexto, de la plenitud de potestad, quedó mostrado en el VI y el VII de esta Parte que la potestad de atar y desatar culpas y penas, y similarmente la de anatematizar o excomulgar a alguno públicamente, no ha sido concedida al sacerdote absolutamente, es decir, con plenitud, sino que ha sido determinada por ley divina, de forma que no pueda, ante Dios, condenar a los inocentes ni absolver a los culpables. Y además, porque la potestad de excomulgar a alguien públicamente y de poner en entredicho a alguna comunidad, debe convenientemente ser determinada para cualquier obispo o sacerdote por la ordenación humana, como lo hemos mostrado en el VI, VII y XXI de esta Parte. Y más; se mostró en el 17º de esta Parte que la misma potestad de instituir, por la imposición de manos, a los ministros eclesiásticos, de enseñar y predicar y administrar los sacramentos a las comunidades de fieles, no les compete a los obispos y sacerdotes con plenitud tal que no esté determinado por ley divina y humana el modo conveniente de obrar conforme a ella.

De los restantes modos, el séptimo y el octavo de la plenitud de potestad, se mostró en el XX, XXI y XXII de esta Parte que no les conviene con plenitud a ningún obispo o sacerdote, sino según la disposición, tanto de la ley divina, como de la humana. Así pues, no pertenece la plenitud de potestad al obispo de los romanos ni a ningún otro sacerdote, en cuanto tal, a no ser que quiera llamar acaso plenitud de potestad a la primacía o principalidad que mostramos en el XXII de esta Parte convenir al dicho obispo y a su iglesia sobre todas las demás, con la autoridad del legislador humano fiel.

[...]

§ 11. Pero no contentos con esto, sino queriendo escalar la cima de los poderes seculares, contra el precepto y consejo de Cristo y los apóstoles se lanzaron a la promulgación de leyes, al margen de las que se dan para el conjunto de los ciudadanos, declarando al clero todo exento de aquéllas, causando así cisma civil y pluralidad de principados supremos, lo que mostramos en el XVII de la Primera Parte, y con averiguada experiencia, ser incompatible con la paz de los hombres. Esta es la raíz y el origen de la peste del reino de Italia, de donde han brotado y brotan todos los escándalos, y permaneciendo la cual, nunca se acabarán las contiendas civiles. Esta potestad, a la que se fue acercando poco a poco y con sinuosa prevaricación

por la costumbre, o mejor, por el abuso, la detentó largo tiempo el obispo de los romanos, y temiendo que le fuera arrebatada por el príncipe justamente, por sus extralimitaciones, con maligna solicitud veta la creación y la promoción de dicho príncipe, y a tanto llegó finalmente su osadía, que afirmó en sus edictos que el príncipe romano estaba atado a él por un *juramento de fidelidad*, como sujeto por jurisdicción coactiva, como lo muestra claramente aquélla de sus fantasías que llaman *decretales*, objeto de burla y menosprecio, y será ello patente a quienquiera que lea el capítulo 7º del *De la sentencia y de la cosa juzgada*².

§ 12. Por no plegarse a esta desatentada temeridad el cristianísimo emperador, hombre de todas las virtudes en todo tiempo, lugar y estado, entre todos los otros príncipes singularmente estimado; Enrique VII de feliz y divina memoria, es declarado trasgresor como *simulador desmemoriado* del juramento prestado en una falsa no menos que temeraria fantasía llamada *decretal*, que lleva por título: *Del juramento*, bien que con más mérito se habría de apellidar de la inicua injusticia y de la ofensa inferida al divino emperador, a sus sucesores, parientes y allegados. Porque como perjuró es declarado infame por los llamados *fundadores de los cánones*. Los que se han empeñado en denigrar su clara memoria, si con palabras o escritos de tales calumniadores pudiera mancillarse.

§ 13. Estas oligárquicas ordenaciones no se atrevieron los obispos de los romanos con sus cardenales a llamar *leyes*, las llamaron *decretales*, bien que sea su intención obligar con ellas a los hombres a la pena para el estado del presente siglo, del modo como pretenden los legisladores humanos. De esta palabra *ley* no se atrevieron a servirse al principio por temor a la resistencia y censura del legislador dicho, porque con ello cometieron un crimen de lesa majestad contra los gobernantes y legisladores. También llamaron a estas ordenaciones al principio *derechos canónicos*, para que, so color del vocablo, impíamente aplicado, se las tenga por más auténticas e impriman en los fieles más respeto y obediencia hacia ellas.

Así poco a poco, y como quien no quiere la cosa, por decirlo todo en una palabra, se han transferido a sí mismos los obispos de los romanos las seis últimas significaciones de plenitud de potestad, y por ellas han cometido en el orden civil multitud de cosas monstruosas contra la ley divina y humana y contra el recto juicio de cualquiera que haga uso de su razón. De algunas de las cuales, de algunas, no todas, hemos, al menos, hecho mención en particular en el capítulo precedente.

[...]

² Cfr. *Corpus Iuris Canonici* (VII Clementinae), l. II, tit. XI, c. 2.

XXX. REFUTACIÓN DE LAS RAZONES ADUCIDAS PARA LO MISMO EN EL MISMO
CAPÍTULO III, Y DE LA TRASLACIÓN DEL IMPERIO ROMANO Y DE CUALQUIER OTRO
PRINCIPADO, EN CUANTO DEBE Y PUEDE HACERSE SEGÚN LA RECTA RAZÓN

[...]

§ 2. Y lo que se aducía además, como razón, que como se compara lo corporal a lo espiritual así el príncipe de lo corporal al que lo es de lo espiritual, habrá que rechazarlo de igual modo que la anterior objeción, pues en ella se apoya como en su raíz. Porque la mayor admite muchos argumentos en contra; a la menor, en la que se dice que lo corporal se ha de someter a lo espiritual, si se entiende someterse como ser más perfecto, se ha de conceder, según las propias significaciones de las palabras, *temporal* y *espiritual*. Pero cuando se añade que el obispo romano es príncipe o juez de las cosas espirituales, si *juez* se toma según la primera significación de juez, como que es quien juzga de estas cosas especulativa y operativamente, es verdadero que el obispo romano y cualquiera otro es y debe ser un tal juez, y de ahí se concluye que es más perfecto que el que juzga con un tal juicio solamente de las cosas corporales, principalmente por la diferencia de las cosas juzgadas. Pero de ahí no se sigue que el juez de las cosas espirituales sea superior al otro en términos de jurisdicción o de juicio coactivo. Porque en ese caso el que se cuida de los animales sería príncipe o juez coactivo del astrólogo o del geómetra, o viceversa, siendo así que ni lo uno ni lo otro es necesario ni verdadero. Pero si se entiende que el obispo romano u otro cualquier obispo es juez de las cosas espirituales según la tercera significación, a saber, coactiva, hay que negarla como manifiestamente falsa, como de la carta de Santiago, 4º lo dedujimos y en el IX de esta Parte. Porque sólo Cristo es de esta forma juez, al que nunca hemos negado ni negamos que deba someterse el juez de este siglo en términos de jurisdicción coactiva para el estado del siglo futuro. De donde el apóstol en la carta a los Efesios, 6º y a los Colosenses, último: *El Señor de ellos y vuestro está en los cielos*¹, cuando no había entonces ningún apóstol o sacerdote de la Nueva Ley fuera de Cristo. Por lo que sólo por este juez serán juzgados con juicio coactivo los jueces de este siglo, y los que desmerecieren serán castigados por el poder coactivo, pero en el otro mundo, conforme a su ley, como quedó claro en el IX de esta Parte. Por la equivocidad, pues, de este vocablo *juez* fallaba el dicho silogismo.

§ 3. Y a aquella argumentación de que como el fin al fin, la ley a la ley, etc., así el juicio al juez, puede negarse tomada en su universalidad. Pero tomada y concedida indeterminadamente con la menor adjunta, puede

¹ *Efes.*, 4, 9; *Colos.*, 4, 1.

concluir acaso por razón de la materia, que el juez coactivo según la ley divina es superior al juez coactivo según la ley humana, lo que hemos concedido antes. Pero si se entiende que el obispo romano o cualquier otro obispo es juez según la ley divina, habrá que distinguir la proposición mirando la equivocidad de este término *juez*, y hay que negarla en el sentido que pretende concluir el objetante, que el romano o cualquier otro obispo es juez coactivo en este siglo o en el futuro según la ley divina.

[...]

§ 5. Y lo que se objetaba después, a saber, el inconveniente de que el vicario singular de Cristo, el obispo romano o cualquier otro sucesor de los apóstoles, deba someterse al juicio coactivo del que es gobernante sólo por la ley humana, hay que decir que no hay ningún inconveniente en que el vicario de alguien se someta a aquel o similar juez, al que el señor del mismo vicario espontáneamente determinó someterse para que se conservara el orden conveniente en este mundo. Porque Cristo, Dios y hombre, se sometió espontáneamente al juicio coactivo de Poncio Pilato, gobernante según la ley humana (vicario del César); así también hicieron los santos apóstoles y a otros mandaron hacer lo mismo según la ley divina, como claramente se demostró por la Escritura y los dichos de los santos y otros doctores en el IV y V de esta Parte, y se repitió en el XXVI. De donde, como *el siervo no es más que su señor ni el apóstol más que el que le envía*², como por Bernardo adujimos de la Escritura en el XXVIII de esta Parte, ningún inconveniente, sino más bien es muy conveniente, más aún, necesario para la quietud de la ciudad o de la comunidad política, que todo obispo y presbítero y clérigo se someta al juicio coactivo de los gobernantes según la ley humana. Y lo contrario de esto es totalmente inconveniente e intolerable, como se demostró en el XVII de la Primera Parte y en ésta en el pasaje arriba citado, confirmado con testimonios eternos.

Todavía, porque el obispo romano o cualquiera otro no es vicario de Cristo o ministro para ejercer todo oficio en este mundo, sino sólo uno determinado, *verbi gratia*, el sacerdocio, en el que, en cuanto tal, no se comprende el juicio coactivo inferior o superior en sí mismo considerado, como lo adujimos de Aristóteles claramente en el IX de esta Parte, párrafo 8³; mas el gobernante, según la ley humana, es vicario y ministro de Dios en cuanto al oficio de gobernar, en el que lo superior y lo sometido se entienden por relación a la potestad coactiva. De donde en la carta a los Romanos, 13^o, sin excluir a nadie, obispo o presbítero, dice: *Toda alma se someta a las potestades supremas*, añadiendo la causa, *porque es ministro*

² Juan, 13, 16.

³ Polit., l. IV, c. 15, 1299 a 16 s.

de Dios⁴. He ahí el vicario de Dios, no cualquiera, sino coactivo para los malos en este mundo. Por donde añade: *Es vengador en su cólera contra el que obra el mal*. Pero obrar mal puede el obispo o el sacerdote, para quienes nunca Cristo ni apóstol alguno designó, de palabra, de obra o con el ejemplo, otro juez, como lo demostramos en esta Parte, en el pasaje arriba citado.

§ 6. A aquella objeción que se proponía por vía de pregunta, a saber, si es conveniente corregir a los gobernantes por medio de la autoridad humana cuando delinquen contra la ley divina o humana, y no parece que puedan ser corregidos convenientemente dado que no tienen, los primeros, al menos, o el primero de ellos, un superior en la vida política, y por ello deben someterse al juicio coactivo de los sacerdotes o de los obispos, hay que decir que el gobernante que delinque contra la ley divina o humana puede y debe ser corregido convenientemente por el ministro eclesiástico, obispo o presbítero, por medio de palabras exhortatorias o increpatorias, siempre modestas según la enseñanza del apóstol, 2ª a Timoteo, 2º y 4º y el comentario del Crisóstomo⁵, que adjunimos en el IX de esta Parte, pero de ningún modo por vía de potestad coactiva, porque ésta en modo alguno pertenece al obispo o al sacerdote, en cuanto tales, sobre nadie en este siglo, como muchas veces lo hemos probado y repetido en lo anteriormente dicho. Pero corregir, de acuerdo con la ley humana, al gobernante por sus abusos contra las leyes, contra la ley, en cuanto ley humana, tal como la definimos en el X (de la Primera Parte), y castigarle, si es preciso, con pena o suplicio temporal, pertenece sólo a la autoridad del legislador humano, o a la de los instituidos por él para eso, como en el XVIII de la Primera Parte pensamos haberlo dejado mostrado suficientemente.

§ 7. La deducción que finalmente se añadía, que en términos de jurisdicción aquél es superior al emperador romano que, con derecho y poder de instituirle y deponerle, *traspasó el imperio de los griegos a los germanos en la persona de Carlo Magno*, y esto es el Papa romano, luego es superior al emperador, y le puede con derecho constituir y deponer, hay que decir que si la mayor se toma de modo indefinido, de ella con la menor no se sigue nada, por no poderse construir un silogismo de una indefinida con una particular. Pero si se toma de un modo universal, de forma que se diga: Todo transmisor del imperio romano de los griegos a los germanos es superior, etc., si no se determina el sujeto, tiene esta proposición verdaderas objeciones. Porque si alguien de hecho y no de derecho transfirió el imperio, con potestad ajena u otorgada a sí por otro para esto, como a un

⁴ Rom., 13, 1-4.

⁵ II Tim., 2, 24; J. Crisost., *De sacerdocio*, II, 3, 4. PG 48, 634 s.

procurador o parecidamente, si así lo hubiere transferido, no por eso sólo tendría jurisdicción superior ni poder de instituir o deponer al príncipe romano. Pero determinada la primera proposición de modo que todo mortal que transfirió o puede transferir el imperio romano de los griegos a los germanos, con autoridad propia, no concedida por otro, es superior en términos de jurisdicción al príncipe romano y le puede justamente instituir y deponer, así se concede. La segunda que le sigue, a saber, que el obispo romano o Papa es quien transfirió el imperio, como se ha dicho, hay que negarla como del todo falsa. Pues la opuesta a ella se demostró en el XV de la Primera Parte, y tanto lo opuesto a ella como a la conclusión deducida de ella se hizo patente por la Escritura y por los dichos de los santos y doctores católicos en el IV y el V de esta Parte, y se repitió en otros muchos lugares, y que así fue sostenido de hecho y sin protesta por los antiguos padres y pastores, los obispos romanos, lo hicimos ver en el XXI de esta Parte deducido de historias aprobadas.

Y en cuanto a lo que se escribe en el VII de ciertos relatos que llaman decretales *Sobre el Juramento* y en una cierta carta del llamado Papa romano al ínclito Luis, Duque de Baviera, elevado a rey de romanos, a saber, que por la sede apostólica, o por el Papa romano, o por él solo, o con el colegio de sus clérigos, *el imperio romano fue transferido (razonable o justamente) de los griegos a los germanos en la persona de Carlo Magno*, dejemos por ahora esta hipótesis, porque de esta traslación, en qué modo fue realizada de hecho, hemos de hablar aparte en un tratado distinto de éste. Demos que el traslado del imperio de los griegos a los germanos se hizo justamente, pero digo que no por la autoridad sólo del Papa romano, o por él con sólo su colegio de clérigos, como dijimos⁶.

§ 8. Y por ello conviene poner atención, a tenor de lo demostrado en el XII, XIII y XV de la Primera Parte, en que para los que miran las cosas según la recta razón, al mismo pertenece la autoridad primera de establecer las leyes humanas, instituir el principado, designar el príncipe, concederle autoridad, y la de mudar, destruir, aumentar o quitar, suspender, corregir, deponer, transferir, revocar y de hacer en torno a estas cosas lo que pareciera bien al que tiene esa autoridad dicha de modo principal, no recibida de otro, y lo expresare por su voluntad. A quién corresponda dicha autoridad lo determinamos en el XII y XV de la Primera Parte. Por lo que dondequiera que se lea o por alguien se diga que fue transferido el imperio, o que cualquier otro principado o príncipe, que se toma por elección, ha sido instituido por el Papa o por otra persona singular, o colegio particular de la provincia o del reino, para que sea verdadero el escrito o la palabra y válida

⁶ *Corpus Iuris canonici. Clementinae.*, 1. II, tít. 9, c. 1. Ver *supra*, DP, II, 23.2.

o justa la institución y semejante traslado, conviene que haya sido hecho, o se haga, con la autoridad del legislador primero en la provincia o provincias, sobre las cuales, por las cuales y para las cuales se hizo o se ha de hacer tal institución o traslado. y por ello, si la transferencia del imperio romano, o la institución de algún emperador se dice o se escribe que fue hecha legítimamente por el Papa romano, o solo, o con solo su colegio de clérigos, para que sea verdadero ese dicho o escrito, es preciso que se entienda que el traslado o la institución dicha ha sido hecha por ellos con la autoridad a ellos concedida para eso por el legislador supremo del imperio romano, de modo inmediato o mediato, o que no ha sido hecha absolutamente por ellos, sino sólo de un cierto modo, como publicada o proclamada, pero con la autoridad antes dicha. Porque acaso transfirió el imperio, o instituyó sobre él, el legislador predicho, el cual luego encomendó a solo el Papa romano, como a la persona más digna de respeto de toda la universalidad humana, o a él con sus sacerdotes, como el colegio de clérigos más venerable, que lo publicaran o proclamaran, no ciertamente como algo necesario para lo hecho o por hacer, sino sólo por razón de solemnidad, porque las transferencias del principado, las instituciones de las leyes y de los gobernantes, igual que de los demás oficios civiles, en cuanto tales, en lo que toca a su vigencia, dependen únicamente de la elección u ordenación del legislador dicho, como hemos demostrado en el XII y XIII de la Primera Parte y no sin utilidad repetido en el XXVI de ésta, párrafo 5, con una declaración.

Así en todo habrá que pensar de la institución del oficio de los príncipes electores del emperador romano, porque no tienen otra autoridad en esto ni les viene de otro, ni les puede ser suspendida ni revocada más que por el dicho supremo legislador humana del imperio romano, Baste haber así recorrido las dudas propuestas en el III y el XXVII de esta Parte y haber así puesto fin a las cuestiones.

PARTE TERCERA

II. DE LA DEDUCCIÓN EXPLÍCITA DE CIERTAS CONCLUSIONES QUE RESULTAN NECESARIAMENTE DE LAS PARTES ANTERIORES. ATENDIENDO A LAS CUALES MÁS FÁCILMENTE PODRÁN ALCANZAR, GOBERNANTES Y SÚBDITOS, EL FIN PRETENDIDO EN ESTE LIBRO

Proponemos la primera de esas conclusiones deducidas.

§ 1. Que sólo es cierta la divina o canónica Escritura, y cualquier interpretación suya que de ella se deduce necesariamente, hecha por el

concilio común de los fieles, y creer en ella, si es debidamente propuesta a alguien, es necesario para alcanzar la bienaventuranza eterna.

§ 2. Definir los sentidos dudosos de la ley divina, sobre todo en los que se llaman artículos de la fe cristiana y en los demás que hay que creer de necesidad para la salud, pertenece sólo al concilio general de los fieles o a su multitud o parte más prevalente, y ningún otro colegio parcial o persona particular, de cualquier condición que sea, tiene autoridad sobre tal definición dicha. La certeza de esto.

§ 3. No se manda en la ley evangélica que se fuerce a nadie, por medio de pena o suplicio, a observar los preceptos de ley divina.

§ 4. Sólo los preceptos de la ley evangélica y los que se siguen de ella con necesidad o los que, según la recta razón, conviene hacer u omitir, han de ser observados para la salud eterna, no todos los de la Antigua Ley.

§ 5. Ningún mortal puede dispensar de los preceptos y prohibiciones de la ley evangélica. Pero prohibir, dentro de lo que está permitido, obligando a la culpa o a la pena para el estado de la vida presente o futura, sólo lo puede el concilio general o el legislador humano fiel, y ningún otro colegio parcial o persona particular de cualquier condición que sea.

§ 6. El legislador humano es sólo la universalidad de los ciudadanos o la parte más prevalente de ella.

§ 7. Las decretales o decretos del pontífice romano y de cualesquiera otros pontífices, colegial o individualmente dictados, sin la concesión del legislador humano, o del concilio general, no obligan a nadie a culpa ni a pena temporal ni espiritual.

§ 8. En las leyes humanas sólo el legislador, u otro con la autoridad de él puede dispensar.

§ 9. Un principado elegido y cualquiera otro oficio depende solamente de la elección de aquel que tiene autoridad para ella, y de ninguna otra confirmación o aprobación.

§ 10. La elección de cualquier principado o de otro oficio que se ha de instituir por elección, principalmente del que tiene fuerza coactiva, depende de la sola voluntad expresada por el legislador.

§ 11. En la ciudad o el reino ha de haber sólo un único principado supremo.

§ 12. A la sola autoridad del gobernante fiel pertenece el designar las personas, su calidad y número, para los oficios de la ciudad, igual que para todo lo demás que comporta el orden civil, según las leyes y las costumbres aprobadas.

§ 13. Ningún gobernante y menos un colegio parcial o persona singular, de cualquier condición que sea, tiene la plenitud del imperio, o de la

potestad sobre los actos privados o civiles de otros, sin la determinación del legislador mortal.

§ 14. Ningún obispo o sacerdote, en cuanto tal, tiene el principado o la jurisdicción coactiva sobre ningún clérigo o laico, aunque sea hereje.

§ 15. Sólo el gobernante, con la autoridad del legislador, tiene jurisdicción coactiva, tanto real como personal, sobre toda persona singular mortal de cualquier condición que sea y sobre el colegio de laicos o de clérigos.

§ 16. No le es lícito a ningún obispo, presbítero o colegio de ellos excomulgar a nadie sin la autoridad del legislador fiel.

§ 17. Todos los obispos están investidos inmediatamente por Cristo de igual autoridad, ni se puede demostrar por la ley divina que, en lo espiritual y en lo temporal, estén entre sí encima o debajo unos de otros.

§ 18. Con autoridad divina, y mediando el consentimiento o concesión del legislador humano fiel, pueden los otros obispos, colegial o individualmente, excomulgar al obispo romano, y ejercer otra autoridad sobre él, y lo mismo a la inversa.

§ 19. A ningún mortal le es lícito dispensar en las uniones conyugales o matrimoniales prohibidas por la ley divina, pero en las prohibidas por la ley humana pertenece solamente a la autoridad del legislador, o del que en virtud de él gobierna.

§ 20. Legitimar a los nacidos de un lecho o matrimonio ilegítimo, de forma que puedan acceder al derecho de la sucesión hereditaria y recibir los otros oficios y beneficios civiles y eclesiásticos, es claro que pertenece sólo al legislador fiel.

§ 21. Promover individuos a los órdenes sagrados eclesiásticos, y juzgar de su capacidad con juicio coactivo, pertenece sólo al legislador fiel, y sin su autoridad no es lícito a ningún sacerdote u obispo promover a nadie.

§ 22. A sólo el gobernante, según las leyes de los fieles, pertenece el señalar el número de iglesias o de templos y de los sacerdotes, diáconos y otros oficios con encargo de servir en ellos.

§ 23. Los cargos eclesiásticos separables, sólo con la autoridad del legislador fiel deben conferirse e igualmente pueden retirarse, y lo mismo los beneficios y las demás cosas establecidas para causas pías.

§ 24. Instituir notarios u otros oficios públicos civiles no pertenece a ningún obispo, en cuanto tal, ni colegial ni individualmente.

§ 25. A ningún obispo, ni colegial ni individualmente, en cuanto tal, corresponde conceder la licencia de enseñar o de ejercer en público un arte

o disciplina, sino que esto pertenece al legislador, al menos si es fiel, o, con su autoridad, al gobernante.

§ 26. Los promovidos al diaconado o al sacerdocio y los demás irrevocablemente consagrados a Dios se han de preferir a los no así consagrados en los oficios y beneficios eclesiásticos.

§ 27. Satisfecha la necesidad de los sacerdotes y de los otros ministros del evangelio y de las cosas que pertenecen al culto divino, puede el legislador usar de los bienes eclesiásticos en todo o en parte para las utilidades comunes o públicas, o para la defensa.

§ 28. Disponer de los bienes temporales que han sido instituidos para las causas pías u obras de misericordia, como lo dejado en testamento para el paso a ultramar, para resistir contra los infieles, o para rescate de los cautivos de ellos, o para la sustentación de los pobres impedidos y semejantes fines, pertenece sólo al gobernante de acuerdo con la determinación del legislador y de la intención del que lo lega o de cualquier otro modo lo otorga.

§ 29. A sólo el legislador fiel pertenece el conceder dispensa a un grupo cualquiera o a un grupo religioso, igual que aprobarlo o instituirlo.

§ 30. Juzgar con juicio coactivo a los herejes y a todos los delincuentes, a los que hay que castigar con pena o suplicio, infligir penas personales y exigir las reales y aplicarlas, pertenece sólo a la autoridad del gobernante, según la disposición del legislador humano.

§ 31. Ninguno sujeto o ligado a otro por juramento lícito puede ser desligado por un obispo o un sacerdote sin causa razonable que habrá de ser estimada por el legislador fiel en juicio según su tercera significación; lo opuesto se enfrenta con la sana doctrina.

§ 32. Instituir un obispo o una iglesia absolutamente metropolitana de todas las otras, y privar o deponer de tal oficio, compete sólo al concilio general de los fieles.

§ 33. Convocar con potestad coactiva, el concilio general o uno parcial de sacerdotes y obispos y demás fieles, pertenece sólo al legislador fiel, o, con su autoridad, al gobernante en las comunidades de fieles, y no tienen fuerza o valor las cosas determinadas en uno congregado de otra manera, ni obligan a nadie a su observancia con pena o culpa temporal o espiritual.

[...]

§ 41. El obispo romano y cualquier otro eclesiástico o ministro del tempo, sólo debe, según la ley divina, ser promovido a un oficio separable, por el legislador o, con su autoridad, por el gobernante, o por el concilio general, y por el mismo ser suspendido o privado, exigiéndolo el delito.

[...]

III. DEL TÍTULO DE ESTE LIBRO

Se llamará este tratado *Defensor de la Paz*, porque en él se tratan y se explican las principales causas por las que se conserva y se da existencia a la paz civil o tranquilidad, y aquéllas también por las cuales la opuesta contienda nace, se impide y se suprime. Por él se da a conocer la autoridad, la causa y la concordancia de las leyes divinas y humanas y de todo principado coactivo, cuáles son las reglas de los actos humanos, en cuya medida conveniente y no impedida consiste la paz o tranquilidad civil.

Por él, tanto el príncipe como el súbdito pueden comprender cuáles son los primeros elementos de cualquier ciudadanía, que hay que observar para la conservación de la paz y de la propia libertad. Porque el primer ciudadano o la primera parte del régimen civil, a saber, el príncipe, ya sea un hombre o muchos, comprenderá, por las verdades escritas en este libro, humanas unas, divinas otras, que a él sólo le compete la autoridad de mandar a la multitud sometida colectiva o individualmente, y de castigar, a cada uno, si es preciso, según las leyes dadas, y de no hacer nada fuera de ellas, sobre todo en lo dificultoso, sin el consentimiento de la multitud sometida, o del legislador, ni provocar a la multitud ni al legislador, porque en la expresa voluntad de éste estriba la autoridad del principado. Y la comunidad sometida, o cualquier individuo de ella puede tener conocimiento, por este libro, de cuál y cuáles gobernantes convenga instituir, y que está obligado a obedecer sólo a los preceptos de la parte gobernante en lo que tienen de coactivo, para el estado y en el estado de la vida presente, y sólo de acuerdo con las leyes promulgadas, y sabrá también en qué cosas están éstas determinadas y en cuáles no, a tenor de lo expuesto en el XV y XVIII de la Primera Parte, y, en cuanto fuere posible, habrá de mirar a que el gobernante, o cualquiera otro de la comunidad, no imponga su arbitrio en el juzgar o hacer alguna otra cosa contra o al margen de las leyes.

Entendidas estas cosas, retenidas en la memoria y guardadas y observadas diligentemente, se salvará el reino y cualquiera otra comunidad civil templada en su ser pacífico y tranquilo, mediante lo cual, los que viven civilmente conseguirán, y sin lo cual de necesidad se privarán, de la suficiencia de vida en este mundo y mal se dispondrán también para la felicidad eterna. Fines los dos, culminación de los deseos humanos, según el uno y el otro mundo, que nos propusimos como meta en las anteriores disertaciones, como algo a todos de suyo manifiesto, añadiendo ahora que, si se hallare en ellas algo determinado, definido o de cualquier otro modo enunciado o escrito con sentido menos católico, eso no ha sido dicho con pertinacia, y lo

sometemos para su corrección o determinación a la autoridad de la iglesia católica, o del concilio general de los fieles cristianos.

Se acabó este *Defensor* (el 24 de junio)
en la festividad del Bautista de 1324.

Loor y gloria a ti, Cristo. ☐

REGULACIÓN DEL FINANCIAMIENTO DE LA ACTIVIDAD POLÍTICA

Comisión Asesora Presidencial para el Fortalecimiento de los Principios de Transparencia y Probidad Pública

Luis Bates Hidalgo (*Presidente*),
Alberto Arenas De Mesa, Francisco Javier Cox Irrarázaval,
Francisco Fernández Fredes, Ángel Flisfisch Fernández,
Eugenio Guzmán Astete, Gonzalo Sánchez García-Huidobro,
Salvador Valdés Prieto, Ángela Vivanco Martínez,
José Zalaquett Daher, Patricio Zapata Larraín y
Rodrigo Medina Jara (*Secretario Ejecutivo*)

El presente informe se terminó de redactar el 5 de mayo de 2003, y tras ser aprobado por la Comisión Asesora Presidencial, fue puesto a disposición del Presidente de la República y otras autoridades.

La elaboración del informe estuvo a cargo de la Subcomisión de Regulación del Financiamiento de la Actividad Política, que se constituyó el 22 de enero de 2003 al interior de la Comisión Asesora Presidencial para el Fortalecimiento de los Principios de Probidad y Transparencia Pública.

Para información sobre otros artículos y documentos dados a conocer por *Estudios Públicos* sobre el tema de la regulación del financiamiento de la actividad política, consultar www.cepchile.cl

1. Introducción

El presente documento contiene un conjunto de propuestas elaboradas por la Subcomisión de Regulación del Financiamiento de la Actividad Política, que es parte de la Comisión Asesora Presidencial de Fortalecimiento de los Principios de Probidad y Transparencia Pública.

Hasta la fecha, la Comisión ha trabajado dividida en dos subcomisiones: la señalada y una segunda sobre Alta Dirección o Gerencia Pública. Si bien el trabajo sobre alta dirección pública está orientado por los objetivos de identificar un sistema o normativa de reclutamiento de directivos que permita asegurar la mantención en el tiempo de un nivel promedio de calidad técnico-profesional lo más cercano a la excelencia, consolidar un carácter meritocrático para el servicio civil, facilitar y promover mayor eficiencia y eficacia de desempeño público, consolidar una cultura moderna de gestión pública y garantizar grados crecientes de igualdad de oportunidades en el acceso a cargos públicos directivos, este cometido también se vincula con la regulación del financiamiento de la actividad política.

En efecto, las prácticas de reclutamiento de directivos hoy imperantes, que privilegian consideraciones político-partidistas, por otra parte en algún grado inevitables, tienden a afectar negativamente tanto los niveles de calidad técnico-profesional históricamente alcanzables, como también los niveles de transparencia, probidad y desempeño ético funcionario al facilitar controles monopólicos que conllevan riesgos de cautiverio, presión e influencias ilegítimas sobre el actuar de los directivos públicos.

Esas prácticas están relacionadas con la ausencia de marcos regulatorios adecuados del financiamiento de la actividad política en sentido estricto: campañas electorales y actividad regular de los partidos. Existen así alimentaciones y retroalimentaciones importantes entre las lógicas de financiamiento de la política hoy imperantes y las prácticas de reclutamiento existentes.

Por consiguiente, los miembros de la subcomisión y por ende los de la Comisión, ven las propuestas que se presentan como complementarias de las que pueda hacer la subcomisión sobre gerencia pública.

2. Valores y principios que deben orientar el marco regulatorio

La Comisión considera que el marco regulatorio que se identifique debe atender a la tutela de diversos valores y orientarse por ciertos principios específicos. Estos valores y principios son los siguientes:

a) Principio de eficacia

La experiencia comparada, que es rica y variada, muestra no sólo que los problemas sobre los que recae esta regulación son complejos, sino también que uno de los talones de Aquiles de los marcos regulatorios es su ineficacia o inocuidad al ser puestos en práctica. La actividad de los distintos actores involucrados —políticos, privados y sociales— tiende a identificar rápidamente oportunidades estratégicas y tácticas para burlar regulaciones. Esta ineficacia, además de tornar inocuo el marco regulatorio, genera frecuentemente escándalos que contribuyen a desprestigiar aún más la actividad política. Por consiguiente, una preocupación permanente debe ser la de construir un marco regulatorio que sea eficaz. No basta con normas imperativas y prohibitivas. Es necesario también identificar incentivos, positivos y negativos, que conduzcan “naturalmente” a los agentes a adecuarse a las regulaciones, y diseñar mecanismos efectivos de fiscalización y castigo.

b) Igualdad de oportunidades

Las modalidades de financiamiento de la actividad política impactan de manera bastante directa en las *chances* de competición política, no sólo por el más que importante peso específico de los recursos financieros en éxitos electorales, sino también al establecer barreras de entrada a la competición, congelando el universo de contendientes. La Comisión estima que el progreso en una razonable igualdad de oportunidades en esta materia es un anhelo mayoritario en la sociedad chilena y ciertamente un factor en la elevación de la calidad de nuestra democracia. El marco regulatorio debe atender a este objetivo.

c) Austeridad

Existe consenso en la sociedad chilena que, respetando la importancia de la actividad política, hay sin embargo presente una tendencia a niveles de gasto excesivo y probablemente ineficiente en las campañas políticas. Ese exceso e ineficiencia en el gasto provoca un creciente rechazo de opinión pública al darse en un contexto de restricciones económicas y evidentes prioridades sociales, contribuyendo a disminuir el prestigio de la actividad y sus protagonistas. El objetivo de introducir un mayor ascetismo

o austeridad, especialmente en las campañas, también debe ser atendido por el marco regulatorio.

d) Prevención de corrupción

La experiencia muestra que en ausencia de regulación existen riesgos muy reales de que se generen prácticas corruptas. Entre ellas hay que destacar las siguientes:

- (i) Comercio voluntario de favores políticos, futuros o presentes, entre candidatos o incumbentes*, por una parte, y privados por la otra;
- (ii) Captura aquiescente de incumbentes, candidatos o partidos por financistas de la actividad;
- (iii) Extorsión o logro coactivo de recursos extraídos a privados por actores políticos a partir de amenazas de represalias.

El marco regulatorio debe procurar dificultar al máximo estas prácticas, de modo de convertirlas en excepciones poco frecuentes. Este objetivo se asocia con el siguiente.

e) Transparencia

Un marco regulatorio que haga transparente el financiamiento es el mejor antídoto a las prácticas corruptas señaladas, pero a la vez posee un valor en sí: La transparencia es parte de la calidad de la democracia.

f) Libertad del ciudadano

Hay aquí dos dimensiones involucradas:

- (i) El ciudadano es libre y autónomo para decidir a quién financia y en qué medida lo financia;
- (ii) El ciudadano debe ser resguardado de posibles represalias al ejercer su derecho a expresarse al financiar actividades políticas.

La primera dimensión apunta a cuestiones como la legitimidad de límites a las donaciones, en tanto la segunda indica posibles excepciones a

* Tomado del sustantivo inglés *incumbent*, que se refiere a la persona que desempeña un cargo oficial. (N. del E.)

la publicidad de aportes o donaciones, exigida por el imperativo de transparencia.

g) Simetría en la información

La situación actual se caracteriza por notables asimetrías de información. Candidatos y partidos se informan sobre las preferencias de los ciudadanos, conocen a quienes los financian y saben de cuánto financiamiento disponen o cuánto han gastado. Inversamente, los ciudadanos a lo más pueden conjeturar sobre disposiciones, motivaciones y preferencias de los actores políticos, ignorando cuánto gastan y en qué lo gastan y desconociendo quién los financia, por lo que les es difícil o imposible acceder a información útil para evaluar su desempeño. El marco regulatorio puede y debe contribuir a tornar más simétrica la relación entre representantes y representados.

h) Colisión y equilibrio de principios y valores

Los problemas involucrados en la regulación del financiamiento de la actividad política implican frecuentemente colisiones entre los valores y principios que definen sus objetivos. Ello obliga a buscar equilibrios que posibiliten sólo la realización parcial de los valores y principios en juego, o bien, a privilegiar un determinado valor o principio, sacrificando otros. Por ejemplo, la austeridad y la igualdad de oportunidades hacen conveniente establecer límites máximos al gasto de campaña, pero la libertad de expresarse mediante aportes o de disponer de lo propio torna dudosa esa restricción, o bien, la transparencia aconseja que toda donación sea pública, pero hay que resguardar a donantes que podrían ser objeto de represalias. Ello es inevitable y en esta presentación se ha procurado hacer explícitos esos equilibrios o esas decisiones.

3. Ámbito de aplicación del sistema regulatorio

La propuesta que se presenta ha sido concebida fundamentalmente para las elecciones de diputados y senadores. No obstante, es nuestra opinión que ella puede también extenderse en lo esencial, probablemente con algunas adaptaciones, a:

a) Elecciones municipales, restringiéndola a la elección de alcaldes, exceptuando la de concejales, y sólo en el caso de las comunas mayores, por ejemplo, el caso de las cuarenta comunas con mayor población.

b) Elecciones presidenciales, tanto en primera como en segunda vuelta.

Adicionalmente, las proposiciones de regulación incluyen también el financiamiento de las actividades de los partidos políticos en períodos entre elecciones o no electorales.

4. Financiamiento privado de campañas

Los problemas que plantea el financiamiento de campañas a través de donaciones privadas constituyen una de las cuestiones centrales de la regulación del financiamiento de la actividad política. Estos problemas ponen en juego casi todos los principios y valores cuya tutela orienta el esfuerzo regulatorio. Así, entre otras cosas:

- a) Hay riesgos de corrupción al imperar la práctica de donaciones reservadas, que es hoy la regla universal.
Las donaciones no registradas públicamente, sólo conocidas por candidato y donante en cuanto a la identidad del donante y la cuantía, abren claras oportunidades para comercio voluntario de favores políticos, futuros o presentes, extracción coactiva de recursos desde privados por candidatos y captura de candidatos por privados. Probablemente, estos riesgos se acentúan en el caso de candidatos incumbentes.
- b) En la situación actual, hay asimetrías de información entre candidatos y electores, un resultado obvio de la práctica de donaciones reservadas.
- c) Adicionalmente, hoy es imposible identificar la cuantía total de aportes hechos por un determinado privado —persona natural o jurídica—, puesto que, además del secreto que envuelve a las donaciones, es posible dividir grandes donaciones en una diversidad de pequeñas donaciones, escondiendo situaciones que equivalen a la virtual captura de una candidatura.

No obstante, al diseñar el marco regulatorio es necesario también ocuparse de tutelar los siguientes valores o bienes en juego:

- a) No restringir ilegítimamente el derecho de todo ciudadano a expresarse políticamente vía donaciones a candidatos.
- b) Resguardar a donantes de represalias al hacerse pública su identidad.

En armonía con lo señalado, las orientaciones básicas de las propuestas de regulación son las siguientes:

- a) Universalidad del principio de transparencia en cuanto al registro de la identidad del donante y la cuantía de la donación: Toda donación debe hacerse vía la autoridad electoral o un mecanismo funcional equivalente, que registra la identidad del donante y la cuantía de su donación.
- b) Universalidad del principio de publicidad de la cuantía de las donaciones recibidas: La ciudadanía puede informarse sobre la cuantía total de los dineros recibidos por cada candidato, cualquiera que sea la modalidad para donar elegida por el donante.
- c) Distinción de dos modalidades para donar: modalidad pública y modalidad confidencial.

Esta distinción obedece, entre otras, a dos razones principales:

- (i) Se ofrece un mecanismo —la modalidad confidencial— que permite resguardar a ciudadanos que temen represalias o perjuicios de hacerse pública su identidad como donante, o que simplemente prefieren permanecer anónimos.
- (ii) Esta modalidad confidencial, en los términos en que se la ha construido, permite también establecer obstáculos más que importantes a la captura de candidatos por pocos donantes importantes, según se explicará posteriormente.

En todo caso, una donación sólo se puede hacer a través de una de las dos modalidades: Se prohíbe cualquiera otra forma de donación.

- d) Atendiendo a que las donaciones a candidatos constituyen un ejercicio de la libertad de expresión política, no se proponen límites a las cantidades que se donen por una misma persona, siempre que esas donaciones se hagan a través de la modalidad pública, puesto que esa modalidad implica publicidad no sólo de los montos sino también de la persona del donante.

Inversamente, se proponen restricciones a la cuantía cuando el donante utiliza la modalidad confidencial. Estas restricciones también se hacen extensivas a los casos en que la donación es imputable a grupos de personas relacionadas que utilizan esa modalidad.

- e) Las propuestas se han construido de manera que, en cualquier caso, haya simetría de información entre candidatos y ciudadanos respecto de la cuantía de las donaciones recibidas por los primeros. Sólo hay asimetría en el caso de la modalidad confidencial, en la que la identidad del donante es conocida sólo por el donante y la autoridad electoral y no así por el candidato y la ciudadanía, pero vale la pena subrayar nuevamente que aún en estos casos la simetría se mantiene respecto de los montos.
- A continuación, se explican las dos modalidades de donación.

4.1. La modalidad pública de donación

Esta modalidad puede ser utilizada por cualquier persona, natural o jurídica, incluyendo los extranjeros residentes en el país que están legalmente habilitados para ejercer el derecho a sufragio, y de acuerdo a la lógica global del marco regulatorio propuesto, es la que forzosamente deben emplear tanto los propios candidatos en términos de los propios recursos gastados en la campaña, como los partidos al aportar recursos al candidato.

Elas deben sujetarse a las siguientes exigencias:

- a) Las donaciones se hacen a través del Servicio Electoral, o algún mecanismo funcionalmente equivalente, que garantizará su registro oportuno y completo.
- La mención de un “mecanismo funcionalmente equivalente” tiene el siguiente sentido. La magnitud y complejidad del marco regulatorio propuesto, particularmente respecto de la operación de las dos modalidades de donación contempladas, supone un reforzamiento importante del Servicio Electoral, que implicaría un mayor gasto público significativo. En el caso de la gestión de las modalidades de operación, los mayores costos se generan estacionalmente lo que hace razonable pensar en la posibilidad de externalizar esta gestión mediante licitación a agentes privados, o subcontratando, sin perjuicio de preservar las atribuciones propias del Servicio, particularmente las de fiscalización. Por eso, se habla de un mecanismo funcionalmente equivalente.
- b) Cada candidato debe hacer públicas periódicamente las donaciones que va recibiendo, informando a la ciudadanía sobre sus montos e identidades de los donantes.

- c) Las donaciones que utilicen esta modalidad no pueden ser inferiores a un cierto límite mínimo, que se propone fijar en 200 UTM (aproximadamente, \$5,8 millones).
- La razón de este límite mínimo reside en proteger a potenciales donantes pequeños del riesgo de extorsión por candidatos inescrupulosos. En efecto, si donaciones bajo ese límite pudieran hacerse mediante la modalidad pública, se abriría la oportunidad para coaccionar a ese tipo de donantes exigiéndoles donaciones públicas, lo que proporcionaría al político inescrupuloso la certeza de que el coaccionado ha cumplido con lo que se le exige. Como se verá, la modalidad confidencial impide al candidato conocer la identidad del donante o procurar identificarlo vía el monto de la donación. De esta manera, la efectividad de las amenazas se atenúa considerablemente.
- d) No obstante, se exceptúan de ese límite inferior las donaciones que los partidos hagan a los candidatos o las transferencias de dinero que el propio candidato haga a su campaña, puesto que esas donaciones, cualquiera sea su monto, siempre deben adoptar la modalidad pública.
- También se exceptúan de ese límite inferior las personas jurídicas con más de cien socios y con ingreso anual superior a 3.000 UTM, incluyendo sociedades anónimas.
- e) No hay límite superior para las donaciones que se hagan mediante esta modalidad.

Creemos que esta modalidad puede ser atractiva para asociaciones gremiales, organizaciones sindicales y similares, casos en que la donación es expresiva de un respaldo público, expresión que es ventajosa tanto para el donante como para el candidato. En la mayoría de los casos, la publicidad de la identidad de donantes cuyas donaciones son de gran magnitud va a encerrar muy probablemente un costo en cuanto a imagen y prestigio tanto para el candidato como para el donante. En este sentido, la ausencia de un límite superior será al menos atenuada, en cuanto a incentivos al gasto, por ese costo probable en que se va a incurrir.

4.2. La modalidad confidencial de donación

La mejor manera de presentar esta modalidad es describiendo esquemáticamente su operación:

El Servicio Electoral abre a cada candidato inscrito una cuenta corriente bancaria que recibe depósitos de donantes, los cuales se transforman en donaciones al transcurrir diez días sin que sean retirados.

Perfeccionada la donación, el Servicio las agrega, aplica el límite al donante previa verificación y contraste de las identidades de los donantes, y emplea una fórmula aleatoria para generar cheques semanales a favor del candidato cuyo monto se hace público, pero sin revelar la identidad de los donantes. De esta manera, es difícil que el candidato infiera quién le donó y la ciudadanía tampoco dispone de esa información. A la vez, el Servicio adopta procedimientos internos para dividir la información, precaviendo que sus funcionarios puedan conocer quién donó, cuánto y a qué candidato.

Esta modalidad está sujeta a las siguientes restricciones y prohibiciones:

- a) La suma de las donaciones originadas en un mismo individuo o grupo económico destinadas a un mismo candidato que se hagan a través de esta modalidad no pueden exceder los siguientes límites :
 - (i) Para candidatos a diputados: 800 UTM;
 - (ii) Para candidatos a senador, 800 UTM, multiplicadas por el número de distritos que comprende la circunscripción;
 - (iii) Para candidatos a Presidente, 400 UTM multiplicadas por 60 (el número de distritos que eligen diputados);
 - (iv) Para candidatos independientes a diputado o senador, los límites respectivos se incrementan en un 20% con el objetivo de compensar la inexistencia en estos casos de la fuente de financiamiento que son los respectivos partidos y promover mayor igualdad de oportunidades en la competición electoral.

Respecto de la magnitud de los límites superiores propuestos, la Comisión ha discutido dos alternativas: una que, privilegiando el principio de transparencia, prefiere límites superiores más bajos (por ejemplo, reduciendo las cifras recién propuestas a la mitad); la otra, privilegiando la prevención del riesgo de captura de candidaturas por unos pocos donantes, expresivos de unos pocos intereses, prefiere cifras de magnitud similar a las arriba propuestas.

- b) Las donaciones confidenciales originadas en un mismo donante o grupo de personas económicamente relacionadas que excedan los límites señalados serán rechazadas por el Servicio Electoral. Para este fin, el Servicio tendrá acceso a las bases de datos del Servicio de Impuestos Internos y otros entes públicos, que permiten identificar grupos económicos o grupos de familiares.
- c) No podrán utilizar esta modalidad confidencial ni los partidos, ni los propios candidatos respecto de sus propias campañas.

Tampoco podrán utilizarla extranjeros —salvo los residentes en Chile legalmente habilitados para sufragar—, empresas extranjeras, empresas chilenas cuyos controladores directos o indirectos sean extranjeros, y personas jurídicas con más de cien socios y con ingresos anuales superiores a 3.000 UTM.

- d) Salvo las prohibiciones expresamente establecidas para cada modalidad y las que se derivan del establecimiento de límites —inferior en un caso, superior en el otro—, cualquiera persona puede utilizar indistintamente ambas modalidades.

4.3. Implicancias del marco regulatorio propuesto

La lógica del marco regulatorio propuesto tiene, al menos, tres implicancias:

- a) Los donantes pequeños tienen que sujetarse a la modalidad confidencial.
- b) Los donantes mayores pueden utilizar la modalidad pública sin límite y la confidencial con un límite superior.
- c) Los partidos y los propios candidatos sólo pueden emplear la modalidad pública.

Como se señaló, los principios orientadores y los valores que se busca tutelar son varios y en muchos escenarios posibles entran en colisión. La lógica de la regulación propuesta ha procurado identificar soluciones equilibradas, que posibiliten al menos no sacrificar totalmente los valores y principios en contradicción.

4.4. El caso de las personas jurídicas con número significativo de socios

Un caso especial es el de los donantes que son personas jurídicas con un número significativo de miembros, asociados o socios no controladores. Por ejemplo, asociaciones gremiales, organizaciones sindicales o sociedades anónimas.

En el contexto general de la defensa y promoción de la transparencia y la probidad, es claro que hay derechos de los miembros no controladores que un marco regulador del financiamiento de la actividad política debería tutelar. Al menos, derechos de propiedad, en cuanto las donaciones a candidatos o partidos son gravosas para ellos, y derechos a la libertad y autono-

mía de la expresión política en cuanto las donaciones que se decidan pueden ser contradictorias con las preferencias de esos miembros.

La tutela de esos derechos lleva a proponer el siguiente régimen especial tanto para las personas jurídicas con más de cien socios o miembros, cuyo ingreso anual sea superior a las 3.000 UTM, como para las sociedades anónimas abiertas:

- a) Establecer la obligación para el gerente general o su equivalente de entregar una vez al año al directorio una relación detallada de las donaciones de carácter político que se han hecho por esa persona jurídica, sus filiales y otras personas jurídicas que ella controle, en Chile y en el extranjero.
- b) Establecer la obligación para cada director de pronunciarse sobre esa relación y de que se registren los pronunciamientos en acta.
- c) Establecer la obligación de consignar anualmente en la memoria el monto y destinatario de las donaciones informadas al directorio.
- d) Prohibirles el empleo de la modalidad confidencial de donación, aun cuando una mayoría de la asamblea de socios acuerde hacerlo vía esa modalidad.
- e) Establecer la obligación de que toda autorización por la asamblea de socios al directorio para efectuar donaciones de carácter político especifique el monto máximo anual que esas donaciones pueden alcanzar. Esa autorización tendrá validez por un plazo máximo de cuatro años.

Se puede presumir que en la práctica estas regulaciones van a significar que la modalidad confidencial de donación va a ser empleada principalmente por personas jurídicas pequeñas —con menos de cien miembros o ingresos inferiores a 3.000 UTM—, o por personas jurídicas grandes con pocos socios y propiedad concentrada.

En el primer caso, que sea casi forzoso emplear la modalidad confidencial parece ventajoso dada la naturaleza de estas asociaciones, puesto que las hace menos vulnerables a presiones y coacciones por candidatos inescrupulosos y en esta modalidad no hay límite inferior.

En el segundo caso, esas personas jurídicas pueden utilizar ambas modalidades. A primera vista, les otorga claras ventajas en términos de la influencia del dinero en la política. No obstante, el uso de la modalidad confidencial está acotado por la existencia de un límite superior al monto total que pueden donar a una campaña y la modalidad impide la identificación cierta del donante por el candidato. Por otra parte, la modalidad pública implica una alta probabilidad de sanción social y deterioro de imagen

corporativa al ser públicos tanto los montos como la identidad del donante. Por consiguiente, la lógica de la regulación va en contra de una práctica de donaciones “dispendiosas” y dificulta la captura de candidatos por este tipo de personas jurídicas.

5. Financiamiento privado de partidos en épocas de campaña

La coexistencia de un marco regulatorio como el propuesto para campañas con una situación que permite la donación reservada, no regulada, a partidos, tal como se practica hoy, tendería a hacer ineficaz el primero. Hay experiencia de otros países que muestra que ello es así. La necesidad de regular la canalización de donaciones a los partidos durante la época de campaña no necesita de mayor fundamentación.

En esta materia, se propone lo siguiente:

- a) Toda donación a un partido realizada dentro de los doce meses anteriores al día de inicio de una campaña o durante el período de campaña, se ajustará a una de las dos modalidades —pública o confidencial— ya descritas, rigiendo en esta materia todas las restricciones y límites que regulan las donaciones a candidatos. Se encuentra prohibida toda donación que no adopte una de las dos modalidades.
- b) Los aportes en dinero que un candidato hace a su propio partido son una donación y deben sujetarse a la modalidad pública de donación. Los partidos podrán imponer a sus candidatos la obligación de transferir al partido un porcentaje de las donaciones que reciban, pero esa transferencia se considera donación y debe sujetarse a la modalidad pública de donación.
- c) En el caso de la modalidad confidencial, el límite máximo para la suma de las donaciones hechas a un mismo partido durante los doce meses anteriores al inicio de la campaña, originadas en una misma persona, natural o jurídica, o en un mismo grupo económico o relacionado, será de 8.000 UTM para el total del período.

Un tema distinto al de las donaciones es el de las cuotas de militantes. Como no son donaciones, pueden ser reservadas, tal como es la práctica hoy. No obstante, ello podría abrir una vía de evasión del marco regulatorio mediante el expediente de cobrar cuotas extraordinarias durante el período, cuotas extraordinarias que en realidad están disfrazando donaciones. Para hacer frente a este riesgo, se propone lo siguiente:

- a) La magnitud de la suma mensual agregada de cuotas de militantes cuyo origen se mantiene en reserva durante el año de campaña no puede exceder al promedio de esa misma suma cuando se consideran los dos años anteriores al de campaña.
- b) Toda cuota extraordinaria que exceda lo que un militante donó en promedio durante los dos años anteriores se entenderá que es una donación que debe sujetarse a las modalidades ya señaladas.

6. Financiamiento privado de partidos fuera de época de campañas

Vale la pena tratar a continuación el financiamiento privado de los partidos fuera de la época de campañas, tema que también se inscribe dentro de los ideales de transparencia y probidad. En esta materia, la propuesta es que las donaciones se sujeten sólo a alguna de las dos modalidades ya explicadas: pública y confidencial.

Respecto de la modalidad pública, sólo podrá utilizarse para donaciones superiores a 200 UTM por donación. La razón reside en dificultar prácticas inescrupulosas de coacción a donantes pequeños.

Si la modalidad es confidencial, a través de cuentas bancarias abiertas por el Servicio Electoral a los partidos, cada fuente individual —persona natural, jurídica o grupo económico o relacionado— está limitada en sus aportes a un 10% del gasto corriente declarado por el partido en el balance del año anterior. Para impedir que este límite se vulnere, la ley debe declarar nulo todo contrato de fideicomiso o análogo en el cual una persona encarga a un tercero pagar periódicamente a un partido sumas que, acumuladas, superan ese 10% del gasto corriente del año anterior. Igualmente, será nulo todo contrato por el que alguien se obligue a pagar periódicamente sumas que excedan ese límite.

Un último tema a tratar en este punto es el del tratamiento tributario de estas donaciones. Aquí, se propone que el tratamiento sea el mismo que el dado a donaciones en época de campaña, descrito en el parágrafo 9 de este informe.

7. Algunas formas de evasión

En la literatura se discuten también otras formas de transferencia de recursos, específicamente no monetarios, que podrían emplearse para vulnerar el principio de transparencia.

Es el caso de las donaciones en especie. A juicio de la Comisión, es perfectamente posible sujetar estas donaciones en especie al régimen ya descrito para las donaciones en dinero. Deberían efectuarse a través del Servicio Electoral, acogidas a algunas de las dos modalidades permitidas, con procedimientos de tasación que permitan estimar el valor de mercado de lo que se dona, u obligando al donante a entregar una estimación de ese valor de mercado, estimación que podría ser objetada o rechazada por el Servicio.

Distinto es el caso de servicios personales prestados gratuitamente durante la campaña, o de “préstamos” de bienes diversos también a título gratuito, cuando estos “préstamos” se hacen por personas naturales propietarias de esos bienes. En estos casos, se trata del ejercicio de un derecho personal de expresión política, difícil de conceptualizar como donación al financiamiento de la campaña en los términos que aquí interesa. No obstante, en el caso de “préstamos” de bienes —uso de vehículos, uso de construcciones o infraestructura, etc.— por parte de personas jurídicas, particularmente personas jurídicas que persiguen fines de lucro, ellos deberían sujetarse al régimen de donaciones ya descrito, de manera similar a las donaciones en especie.

Otra hipótesis que ha preocupado en este contexto es la de la figura de intermediarios que agregan donaciones de privados para transferirlas a candidatos. El marco regulatorio que se propone torna inocua esta figura, salvo en cuanto a la eventualidad de coacción de donantes pequeños (por debajo de 200 UTM). En efecto, el empleo de la modalidad confidencial, al implicar un límite superior para un mismo donante, estimula la dispersión de donantes, y no su concentración. En la modalidad pública, para ciertos donantes podría ser conveniente guardar anonimato, escondiéndose tras un intermediario, pero como la identidad del donante es pública y este rol seguramente sería asumido por un tipo de persona ubicado en esa zona gris entre el delito y el respeto a la ley, está el claro desincentivo derivado del hecho de que ese intermediario va a tener que justificar ingresos que respalden donaciones en este caso no menores.

El riesgo se plantea con donantes menores de 200 UTM. Dada la cuantía, la vía confidencial es obligatoria, lo que se justifica porque ello lo resguarda frente a posibilidades de extorsión en razón de que en esta modalidad la identidad del donante no se conoce. La coacción que en este caso se dificulta para un candidato, puede ejercerla el intermediario, con la certeza de que es o no efectiva (el dinero se le entrega a él).

Aparte del argumento ya señalado —la necesidad de justificar ingresos—, no se divisan otras modalidades de prevención. Quizás una pueda

consistir en restringir el derecho de una persona a hacer una donación confidencial sólo una vez durante la campaña. Para usar la modalidad, el intermediario tendría que acumular donaciones extraídas, lo que lo tornaría más vulnerable.

Finalmente, se ha identificado el tema de lo que podría llamarse “campañas laterales” de apoyo llevadas a cabo por ciudadanos no insertos en la campaña misma, o la actividad durante las campañas de organizaciones o grupos que promueven temas específicos, positiva o negativamente, cuya naturaleza hace plausible que incidan en la competición electoral.

Por ejemplo, reales campañas llamando a votar o no votar por candidatos que estén o no estén en determinadas posiciones respecto de esos temas. Si bien es imposible regular estas actividades en cuanto al contenido de la publicidad que llevan a cabo, se propone tipificarlas como fraude a la ley o abuso de derecho en cuanto personas naturales o jurídicas que no son candidatos vulneran la lealtad y el juego limpio (*fairness*) en la competición electoral, valores que el sistema electoral público busca actualizar, a través de un gasto publicitario sustancial que distorsiona el contexto propio de una campaña.

Estas conductas serían objeto de una sanción pecuniaria, la acción se limitaría a candidatos inscritos en distritos donde esa publicidad se efectuó y se podría deducir sólo hasta diez días después del término de la campaña, sería competente el Tribunal Calificador de Elecciones, quien resolvería sumariamente y sin recurso contra esa resolución.

8. Límites al gasto de campaña

Con el objetivo de equilibrar los diversos valores y principios en juego, la presente propuesta es contraria a establecer un límite imperativo al gasto de campaña, prohibiendo a los candidatos traspasarlo, bajo sanciones.

En cambio, a partir de un criterio que creemos más realista y eficaz, se ha optado por desincentivar el crecimiento del gasto mediante:

- a) Incentivos indirectos, obstáculos y regulaciones voluntarias entre candidatos;
- b) La figura de un límite opcional o sugerido, que el candidato puede o no aceptar, con costos claros en el caso de no aceptación; y
- c) La figura de un límite sugerido u opcional para el gasto total o agregado de campaña de los partidos, límite que pueden o no aceptar, pero también con costos claros cuando no se lo acepta.

8.1. Medidas indirectas, obstáculos y regulaciones voluntarias entre candidatos

La medida indirecta sugerida reside en acortar la duración de la campaña. Hoy la fecha de inscripción de las candidaturas es 150 días antes de las elecciones. Se propone acortar ese plazo a 90 días. Sin perjuicio de que hubo acuerdo general en este sentido, se manifestó por uno de los comisionados la prevención en el sentido de que ello también puede constituir una barrera de entrada a los candidatos nuevos, particularmente porque ellos requieren no sólo de mayores recursos financieros (razón por la cual debe tenerse cuidado con los límites) sino que también de tiempo. En consecuencia, reducir el período puede tener efectos negativos, pudiendo operar como un mecanismo de eliminación de competencia por parte de los incumbentes.

Por otra parte, el sistema de regulación de donaciones a campañas y partidos ya descrito ciertamente posee el efecto de reducir las donaciones y por consiguiente reducir el gasto.

Se propone, también, introducir en el marco regulatorio normas que fomenten o promuevan “acuerdos de campaña”, de carácter público, entre candidatos que compiten entre sí, sea dentro de un pacto o entre pactos, para regular la distribución física y protección de propaganda callejera y otras materias que estimen relevantes. Estos acuerdos se sujetarían al siguiente marco:

- a) Para fomentar y prestigiar este tipo de acuerdos, se autorizaría el apoyo gratuito de Carabineros para detectar incumplimientos e identificar sus responsables, y los Juzgados de Policía Local serían competentes para conocer de esos incumplimientos y aplicar las multas previstas en ellos. Así, habría *enforcement* público de estos acuerdos.
- b) Los acuerdos deben ser escritos y públicos, y entregados al Servicio Electoral antes de los diez días después del vencimiento del plazo para las inscripciones de candidatos.
- c) Los acuerdos deben prever multas en el caso de incumplimiento que sean superiores a las multas mínimas fijadas por la ley electoral.
- d) Los acuerdos se anulan si se emplean para perjudicar a otros candidatos que no sean parte del acuerdo.
- e) Los acuerdos son nulos si llevan a desinformar al electorado, por ejemplo, si impiden que otros candidatos participen en debates previstos en el acuerdo.

- f) Todo candidato, incluido o no en el acuerdo, tiene derecho a reclamar del incumplimiento del acuerdo ante el Juez de Policía Local.

8.2. Límite opcional al gasto de campaña de candidatos

El sistema de regulación que se propone contempla un límite legal opcional al gasto de campaña que hace un candidato y reconoce el derecho a dos o más candidatos que compiten entre sí a llegar acuerdos que, entre otras cosas, establezcan límites al gasto, por sobre o por debajo de ese límite legal opcional.

Si bien el límite legal es opcional, el marco regulatorio contempla también incentivos, tanto para los candidatos como para los partidos, que operan en el sentido que el gasto no sobrepase ese límite.

El problema central aquí reside en las dificultades de control, verificación y fiscalización del gasto, de modo de determinar si los límites aceptados se han cumplido o no, dificultad igualmente presente en un sistema que establezca un límite legal obligatorio. La experiencia comparada prueba fehacientemente la realidad de esas dificultades.

Es por eso que, sin perjuicio de las atribuciones propias del Servicio Electoral, la propuesta pone énfasis en el fomento y promoción de formas de autorregulación por los propios candidatos, de modo de hacerlos responsables e incentivarlos en el desempeño de esas funciones de control y verificación. Más específicamente, la propuesta contempla:

- a) Un límite recomendado por la ley al gasto total hecho por un candidato en una campaña. Ese límite podría ser de 1.000 UF más 0,04 UF por cada elector en el distrito hasta 300 mil electores, más 0,02 UF por cada elector en exceso a 300 mil electores.
Para un distrito de 100 mil electores, el límite ascendería a 5.000 UF, 33% por debajo del promedio estimado por algunos para campañas en ese tipo de distritos.
En el caso de elecciones senatoriales se podría aplicar la misma regla considerando la totalidad de electores de la circunscripción, y habría que discutir y afinar un límite razonable en el caso de elecciones presidenciales.
- b) Dos o más candidatos pueden presentar al Servicio Electoral, dentro de los cinco días después de la fecha de inscripciones de candidatos, un acuerdo de campaña que establezca límites parciales y globales al gasto total en que incurrirá cada una de las partes del acuerdo, acuerdo que será público al aceptarlo el Servicio Electoral.

- c) Si el acuerdo establece límites al gasto superiores al sugerido por la ley, los candidatos sufrirán una reducción del máximo del subsidio para franja de radio, que es de \$29,3 millones, reducción que se hará efectiva a una tasa constante. Por ejemplo, una tasa que reduzca el máximo señalado en 20% por cada 0,01 UF por elector de exceso por sobre el límite recomendado por la ley para ese distrito. Adicionalmente, si un candidato firma un acuerdo de campaña que establece límites tales que reducen a cero el subsidio para franja de radio, el partido a que pertenece pierde el derecho a permutar parte de la franja de televisión abierta, derecho que se propone más adelante.
- d) Un candidato que no es parte de ningún acuerdo puede presentar un compromiso al Servicio Electoral que, entre otros requisitos, establezca límite a su gasto de campaña, o acogerse al texto de acuerdo al que han llegado otros candidatos. A este candidato se le aplicarán las mismas reglas ya señaladas: recibirá el máximo de subsidio para franja de radio si el límite es el recomendado por la ley o inferior, el máximo de subsidio se reducirá a una tasa constante en la medida en que el límite va superando el legal, y su partido perderá el derecho a permutar parte de la franja televisiva si el subsidio se reduce a cero.
- d) Se entenderá que el candidato que no sea parte de ningún acuerdo, no presente compromiso o no se acoja a un acuerdo de otros candidatos, no acepta límite al gasto de campaña. Este candidato no tiene derecho a subsidio para franja radial y su partido pierde el derecho a permuta ya señalado.
- e) Si la suma total de donaciones percibidas por un candidato al término de la campaña, a través de cualquiera de las modalidades permitidas, excede al límite fijado en un acuerdo, compromiso o adhesión a un acuerdo, pero el gasto se ha sujetado a ese límite, el excedente donado irá en beneficio fiscal.
- f) Todo acuerdo de campaña, compromiso o adhesión a un acuerdo dará derecho a subsidio si, además de establecer límite al gasto, se sujeta a lo siguiente:
- (i) Establece un listado de gastos permitidos.
 - (ii) Define mecanismos de fiscalización y verificación para cada tipo de gastos y establece garantías del pago de las multas.
 - (iii) Extiende el derecho a reclamo por su incumplimiento a todos los candidatos inscritos en el distrito, incluidos los que no son parte del acuerdo.

- (iv) Establece en caso de transgresiones comprobadas a los límites comprometidos una multa equivalente al triple del exceso, en beneficio del Servicio Electoral.
- (v) Si el candidato es parlamentario, debe incluir entre los gastos que se suman para determinar el gasto global el 100% de la asignación parlamentaria no destinada a financiar estudios de utilidad legislativa, percibida en los últimos doce meses anteriores a la elección.

Adicionalmente, la ley tiene que establecer multas que castiguen a los candidatos denunciados de incumplimientos de acuerdos, compromisos o adhesiones a compromisos que prueben ser falsos y el beneficiario de ellas.

8.3. Límites al gasto de campaña por los partidos

Una situación en la que por una parte los candidatos, incentivados por subsidios fiscales, aceptan voluntariamente limitar sus gastos de campaña, y en cambio, por otra parte, los partidos pueden gastar ilimitadamente, carece de sentido. Sería una situación semejante a lo experimentado en el Reino Unido, donde hay límites estrictos al gasto en el distrito, pero al no existir límites a lo que gastan los partidos, el límite al gasto distrital pierde toda efectividad.

Por otra parte, hay un conjunto de servicios comunes que los partidos prestan a sus candidatos y que requieren de financiamiento del partido y el correspondiente gasto.

El objetivo no es prohibir el gasto de los partidos. Tampoco se ha optado en esta propuesta por establecer un máximo impuesto por la ley. Al igual que en el caso de los candidatos, se ha preferido un marco que genere incentivos claros para limitar el gasto. La propuesta es la siguiente:

- a) Si todos los candidatos pertenecientes a un partido han accedido al subsidio de franja radial, ese partido no podrá gastar durante la campaña más que la cantidad resultante de sumar el 30% de todos los límites recomendados por la ley a los diversos candidatos del partido, más una cantidad fija de 8.000 UTM.
- b) La cantidad fija se explica por el objetivo de proteger a los partidos nuevos o más pequeños, que casi con certeza entran a la competición con desventajas y un número menor de candidatos.

- c) Si de los candidatos pertenecientes a un partido algunos han accedido sólo parcialmente al subsidio de franja de radio, o no han accedido por haber optado por gastar sin límite, el partido puede elegir:
 - (i) Gastar sin límite, en cuyo caso sus candidatos que han accedido al subsidio de franja radial lo pierden y tanto el partido como sus candidatos no pueden recibir donaciones a través de la modalidad confidencial.
La Comisión discutió la posibilidad de “castigar” en esta hipótesis al partido privándolo del derecho a financiamiento público en el siguiente período entre campañas. La idea parece algo draconiana. Como alternativa, se lo podría privar de un porcentaje de ese subsidio.
 - (ii) Acogerse al límite al gasto total del partido ya indicado.
- d) Los partidos llevarán una contabilidad pública y sólo podrán reclamar por el incumplimiento del límite, cuando corresponda, los otros partidos legalmente existentes.

Todo exceso de gasto comprobado, por sobre el límite, implicará una multa equivalente al triple de ese exceso.

Si las multas aplicadas en casos de excesos superan el 20% del límite, el partido perderá su derecho a permutar parte de la franja televisiva en la próxima elección.

9. Tributación de las donaciones durante la campaña

Todo marco tributario genera incentivos privados y afecta recursos públicos. En este sentido, un tratamiento especial para las donaciones a candidatos y partidos durante la campaña se encuentra en una suerte de camino intermedio entre financiamiento público y financiamiento privado.

Por otra parte, considerando que la promoción de igualdad de oportunidades es un objetivo del marco regulatorio, no sería del caso subsidiar las donaciones hechas por ciudadanos pudientes, permitiendo que ellas se descuenten como gastos de sus empresas, lo que se refuerza por la consideración adicional de que esas donaciones no son necesarias para generar la renta de esas empresas.

En virtud de estas razones se propone un régimen de excepción con las siguientes características:

- a) Toda donación en dinero, sea por la modalidad pública o la confidencial, queda libre de impuesto a la renta, tanto personal como de empresas.
Para ello, se considerarán un gasto necesario para producir la renta, pero no podrán contabilizarse como gasto para el efecto de determinar la utilidad imponible.
- b) Todas las donaciones a candidatos o partidos están libres del trámite de insinuación.
- c) Todas las donaciones están sujetas al impuesto a las donaciones, lo que permite que tributen más las campañas y partidos que reciben mayor financiamiento privado.
En el caso de las donaciones a través de la modalidad pública, la base del impuesto será cada donación individual, y el candidato o partido retendrá un 5% que depositará de inmediato en el Servicio Electoral, suma que se liquidará al término de la campaña y una vez determinada la cantidad donada durante la campaña.
Si la modalidad es la confidencial, la base del impuesto será el total de lo donado por un mismo grupo económico o persona durante toda la campaña. En este caso, el Servicio hará una retención de 1%.
- d) Las donaciones en especie, bienes, servicios y patrocinios originadas en un mismo donante que sea persona jurídica están sujetas al impuesto a la renta y no podrán ser consideradas como necesarias para producir la renta del donante, por lo que los socios la registrarán como retiro y la declararán en su impuesto global complementario. Si la persona es una sociedad anónima, deberá enterar el 35% de lo donado en calidad de retención.

10. Financiamiento público de la actividad política

No creemos necesario argumentar sobre la necesidad de que existan formas de financiamiento público a la actividad política. Aceptando que ello es necesario y valioso, la propuesta incluye tres modalidades de financiamiento público: a candidatos, a partidos en época de campaña y a partidos fuera de la época de campaña.

10.1. Subsidio a candidatos

No parece razonable subsidiar candidatos de manera proporcional a los votos que obtienen porque esa clase de financiamiento crea más venta-

jas para los incumbentes y afecta negativamente a la igualdad de oportunidades.

Como modalidad alternativa se propone una franja de radio, consistente en un subsidio de un 40% a la suma del valor de las facturas por publicidad que medios de comunicación locales presenten al Servicio Electoral, más gastos de correo impreso y electrónico, más gastos de producción y servicio de páginas web, hasta un máximo de 1.000 UTM.

Este subsidio tiene la ventaja de no ser un bono de suma fija, modalidad que puede estimular la proliferación de candidatos independientes con escaso apoyo, pues mientras el candidato no genere facturas por publicidad y por gastos de la clase indicada por un total de 2.500 UTM no habrá alcanzado el máximo, y sólo podrá acercarse a él en la medida en que logre recursos vía donaciones de privados o desde su partido.

Este subsidio se otorgará a las cinco primeras pluralidades del distrito o circunscripción, número que equilibra adecuadamente la necesidad de crear oportunidades a partidos más pequeños, nuevos o a independientes, por una parte, y obstaculiza la proliferación de candidatos independientes menores.

El subsidio operará de la siguiente manera:

(i) El Servicio Electoral sólo lo otorgará a personas, empresas o contratistas inscritos en un registro ad hoc al menos tres meses antes del inicio de la campaña.

(ii) Esas entidades (contratistas) tienen que acreditar su existencia y cierta circulación o actividad mínima durante los dos años anteriores a la campaña.

(iii) Ellas deben declarar, al inscribirse, un máximo total de unidades físicas de servicios que están dispuestas a prestar al conjunto de los candidatos inscritos para efectos de solicitar subsidios, distinguiendo entre tipos de servicios.

(iv) Deben comprometerse a garantizar a cada candidato inscrito el acceso a un número de unidades físicas de servicio igual al 50% del máximo de 1.000 UTM, dividido por el número de candidatos.

(v) Deben registrar en el Servicio Electoral un tarifado detallado de publicidad política para épocas de campaña.

(vi) Registrar el compromiso escrito y público de aplicar ese tarifado de manera uniforme a todos los candidatos, bajo la sanción de perder el acceso a subsidios.

El Servicio Electoral podrá contratar entes independientes de muestreo de publicidad para verificar la correspondencia entre la publicidad

efectiva, las facturas presentadas y el tarifado comprometido, pudiendo negar el pago en caso de incumplimiento y dudas fundadas.

(vii) Si un contratista se niega a prestar servicios a un candidato, pierde el acceso al subsidio, y si lo hace hacia una etapa avanzada de la campaña el Servicio Electoral puede retirar parte del subsidio.

(viii) Los contratistas podrán pedir garantías a los candidatos, lo que se justifica por la incertidumbre sobre si tendrá o no derecho a subsidio, lo que está condicionado por la eventualidad de que alcance o no a estar dentro de las cinco primeras pluralidades, pero no podrán hacerlo por más del 40% del tarifado inscrito en el Servicio Electoral.

(ix) El contratista entregará al Servicio Electoral las facturas causantes de subsidio dentro de los tres días hábiles después de la prestación del servicio: el Servicio Electoral las registrará por orden de llegada, semanalmente hará público el total de los subsidios solicitados a nombre de cada candidato y avisará a los contratistas en cuanto un candidato llegue al 90%, al 95% y al 100% del subsidio máximo.

Cuando un candidato llegue al 95% del máximo, los contratistas no están obligados a garantizarle el acceso al número mínimo de unidades físicas de servicio comprometidas.

10.2. Subsidio a los partidos en épocas de campaña

La Comisión se pronunció por rechazar modalidades de subsidio en dinero a los partidos en épocas de campaña porque afectan de manera negativa significativa a los independientes, atenúan la igualdad de oportunidades en la competición electoral y generan el riesgo de que la cuantía de estos subsidios se escale por el interés de incumbentes y partidos en ellos.

Como alternativa a subsidios en dinero se propone un mecanismo consistente en el derecho de los partidos a permutar con los canales emisores hasta el 50% del tiempo asignado en la franja de televisión abierta por espacios o escenarios locales gratuitos de publicidad e información a votantes, que hayan sido certificados por el Servicio Electoral. En el caso de candidatos independientes podrían permutar el 100% del tiempo asignado.

Estas permutas serían públicas, convenidas entre un partido y el conjunto de los canales a través del Consejo Nacional de Televisión, tendrían que acordarse hasta 60 días antes de la elección y sólo podrían recaer en espacios y escenarios donde la función de información al votante se compra en igual o mejor forma que en la franja televisiva, excluyéndose así permutas de tiempo por dinero u otros bienes de menor valor expresivo.

Dos miembros de la subcomisión, la Sra. Ángela Vivanco y el Sr. Francisco Fernández, han hecho la siguiente prevención frente a esta propuesta, que se consigna a continuación:

Nos merece reservas la fórmula de permuta que se propone, pues la franja de propaganda gratuita constituye jurídicamente una carga u obligación de derecho público sobre los concesionarios de los canales, que no tiene estimación pecuniaria como tal, y, por otra parte, configura a favor de los partidos políticos un verdadero derecho de orden público establecido no sólo en su interés sino en el de la ciudadanía como destinataria de la información política por ese específico medio. En consecuencia, dado que no estamos en presencia de un bien disponible por sus aparentes titulares (los partidos políticos), mal pueden éstos negociarlo por la vía de una permuta o cualquier otro tipo de contrato.

10.3. Subsidio a partidos fuera de épocas de campaña

La Comisión estimó que hay modalidades de subsidio a partidos fuera de las épocas de campaña que son positivas en cuanto pueden asociarse a incentivos para lograr mayor transparencia en la actividad política y, en general, facilitar desempeños con mejores connotaciones éticas.

Se proponen las siguientes modalidades de subsidio:

(i) Un subsidio a las cuotas de militantes inscritos, que contribuyan regularmente, que estén al día y que hayan votado en las elecciones internas. El subsidio asciende a un 30% de la cuota pagada con un máximo de 3 UTM por militante, salvo las cuotas mayores en que la tasa es de un 10%.

(ii) Subsidio para gastos regulares de funcionamiento de los partidos, incluyendo dietas para sus autoridades, consistente en que el Servicio Electoral pague, por cuenta del partido, hasta el 50% de cada gasto de ese tipo, debidamente itemizado.

Si bien es sabido que este subsidio es relativamente impopular, se estima necesario para avanzar en mayor transparencia y probidad, siempre que se sujete al siguiente régimen:

- a) Para recibir el subsidio, un partido debe hacer pública, periódicamente, la totalidad de sus gastos, con la exigencia de un desglose igualmente total de ellos.
- b) La norma legal que fija el porcentaje en que se subsidian los gastos posee un quórum extraordinario más alto que las leyes ordinarias y el mensaje debe ser firmado por el Ministro de Hacienda.

- c) El subsidio sólo podrá solicitarse para ítemes incluidos en una lista de gastos permitidos, que excluya la compra de bienes raíces, regalos del partido a terceros y otros que escapen al concepto de funcionamiento regular.

El subsidio se calculará respecto del precio de referencia que establezca el Servicio Electoral para el ítem en cuestión, precios de referencia que se determinarán en función de los precios observados en las adquisiciones estatales y municipales durante los últimos años para ítemes análogos. Esta lista de precios de referencia será pública. En el caso de las dietas pagadas a dirigentes y remuneraciones a funcionarios del partido, las cantidades de referencia se determinarán por la Escala Única de Sueldos y sus suplementos.

- d) El subsidio tendrá un límite global anual para cada partido, calculado en función del número de votos obtenidos por los candidatos del partido en la última elección y del número de escaños que tenga en el Congreso.

Se sugiere un tope mensual de 0,003 UF por voto obtenido, más una suma menor en función de los escaños que el partido detenta.

- e) Todo gasto que se ha acogido al subsidio debe ser aprobado por el consejo general del partido. Si el consejo general objeta gastos subsidiados, el partido debe devolver el subsidio.
- f) El subsidio podrá también impetrarse durante los períodos de campaña, siempre que se trate de gastos para funcionamiento regular, totalmente correspondientes con los ítemes en los que el partido ha venido gastando en los doce meses previos al período de campaña.
- g) Un partido podrá acogerse al subsidio sólo si sus estatutos contienen normas claras que exijan a sus autoridades superiores que no sean parlamentarios la publicidad de declaración de patrimonio e intereses y la prohibición de prestar servicios a honorarios a consorcios de grandes donantes, nacionales o extranjeros y de otros comportamientos que impliquen conflictos de intereses dada la naturaleza de un partido político.

Esas normas deben ser públicas y el partido tendría que mantener actualizada y a disposición del público la documentación que acredite el cumplimiento continuo de ellas. Adicionalmente, la ley tendría que establecer las sanciones mínimas que los estatutos deberían contemplar para el incumplimiento de esas exigencias y prohibiciones.

La Comisión consideró también proponer que esta exigencia se haga extensiva como condición al derecho a todos los subsidios públicos, incluyendo la franja de televisión abierta.

Obviamente, hay la opción de no contemplar esas normas en los estatutos del partido, perdiendo el derecho al subsidio.

Dos de los miembros de la subcomisión, los Señores Salvador Valdés y Eugenio Guzmán, no recomiendan el subsidio descrito en términos del siguiente voto de minoría:

El subsidio público al funcionamiento regular de los partidos se justifica, en nuestra opinión, por el hecho de que los partidos proveen bienes públicos valiosos que consisten en canalizar la deliberación sobre las prioridades públicas y los métodos de acción del Estado. Se justifica entonces que el Estado financie parcialmente estas actividades, creando también incentivos para que fluya el financiamiento privado que pueda atraer esta legítima y necesaria actividad.

Sin embargo, la modalidad específica de subsidio que se propone no parece dar garantías de eficiencia, porque no genera una traba efectiva a legislaciones posteriores que eleven el porcentaje de subsidio desde 50% a 60% y así hasta llegar a un 100%. Exigir “super mayorías” no alivia mucho este problema porque la experiencia internacional sugiere que en esta materia todos los parlamentarios y también el Poder Ejecutivo tienden a ponerse de acuerdo con facilidad en el futuro.

La publicidad de todos los gastos de los partidos es un paliativo valioso, porque la opinión pública podrá informarse de eventuales gastos excesivos.

Sin embargo, tenemos dudas sobre la capacidad de este paliativo de operar con efectividad, lo que nos lleva a abstenernos de recomendar la adopción del subsidio propuesto.

iii) Un subsidio a actividades de capacitación política realizadas por los partidos, similar al que administra el SENCE, restringido a 30 cursos anuales por distrito electoral.

Este subsidio operará a partir de una comisión o jurado regional que seleccionará las mejores actividades de capacitación política propuestas sobre la base de criterios públicos y objetivos, integrado de forma de garantizar la independencia de los partidos e instituciones interesadas en el subsidio, con la posibilidad de que otras instituciones diferentes de los partidos puedan postular para tutelar los derechos de los independientes.

Se contemplarán diversas exigencias: cantidad mínima de alumnos por curso, número mínimo de horas por actividad, prohibición de que los mismos alumnos repitan cursos o asistan a más de 3 cursos al año, fiscalización estricta del Servicio Electoral, publicidad de los cursos, derecho de los partidos a objetar una actividad subsidiada ante el Servicio.

Se sugiere un monto de subsidio de 2 UTM por hora.

11. Reforzamiento y autonomía del Servicio Electoral

La propuesta presentada obviamente implica un notable reforzamiento del Servicio Electoral, por una parte en cuanto a un mayor presupuesto y planta acordes con las funciones que debería cumplir en el marco regulatorio propuesto, y por otra en cuanto a atribuciones que les permitan dotar a sus inspectores de la eficacia de que gozan, por ejemplo, los inspectores del Servicio de Impuestos Internos. Como se indicó anteriormente, el mayor gasto implicado, y considerando el carácter estacional de muchas de las actividades previstas para el Servicio, se podría afrontar recurriendo a externalizaciones y subcontrataciones para varias de esas actividades con el fin de lograr mayor eficiencia y ahorro en costos. Ello debería estar expresamente autorizado por la ley.

Adicionalmente, el Servicio debería gozar de una mayor autonomía que la que tiene hoy, mediante, por ejemplo, la creación de un Consejo de tres personas que duren 8 años en el cargo, designadas una a una cada cuatro años por el Presidente de la República, con acuerdo del Senado.

ANEXO

Estimación del impacto fiscal de la propuesta:

El impacto fiscal es la diferencia que se produce en las cuentas fiscales al implementar esta propuesta, en comparación a la situación actual. Consideramos los siguientes impactos:

1. Eliminación de la práctica de mostrar el financiamiento privado como gasto de las empresas sujetas al impuesto de primera categoría. Según estudios empíricos, el costo agregado de la campaña parlamentaria alcanzó en 1997 a 2,3 millones de UF, sin considerar el valor de la franja de TV abierta, y por lo tanto financiado en su totalidad en forma privada. Suponemos que el 30% del gasto en campaña se financió con donaciones en bienes y servicios y el resto en dinero, y suponemos que el 80% de las donaciones en dinero fueron mostradas como gastos de empresas, eludiendo un impuesto de 16% en primera categoría (retención) y un impuesto de 35% a los dueños (tasa final). Esto ocurre cada cuatro años. Suponiendo que la adopción de esta propuesta permite eliminar esta práctica, la propuesta genera un mayor ingreso fiscal anual de:

$$T1 = 0,7 \times 2.300.000 \times 0,80 \times 0,35 \times 16.783,6 \times (1/4) = 1.891 \text{ millones de } \$/\text{año}.$$

2. Ingreso por impuesto a las donaciones y herencias, que hoy no se aplica. Considerando que el financiamiento privado en dinero se reduce a la mitad del observado en 1997, debido a las demás propuestas, y queda en 0,7 millón de UF, y que la tasa media de este impuesto alcanza a 1,5%, se obtiene un mayor ingreso fiscal anual de:

$$T2 = 0,015 \times 700.000 \times 16.783,6 \times (1/4) = 44 \text{ millones de \$/año.}$$

3. La franja de radio propuesta ofrece un piso económico a las 5 campañas con más apoyo. Considerando 60 distritos de diputado y 20 circunscripciones de senadores integradas por un promedio de 3 distritos cada una, 5 franjas de radio por distrito, un tope a cada franja de \$ 30 millones para diputados, y un grado de utilización de 50% en promedio (recordar que el candidato debe financiar el 60% de cada factura con fuentes privadas y que otras propuestas reducen las fuentes privadas de fondos), y que las campañas son cada 4 años, el costo total anual es:

$$G1 = (60 + 20 \times 3) \times 5 \times \$30 \times 0,50 \times (1/4) = 2.250 \text{ millones al año.}$$

(Esto equivale a \$9.000 millones por campaña parlamentaria)

No está clara la justificación de una eventual franja de radio para elecciones municipales, pues en ellas los escaños son mucho más numerosos, lo que asegura la representación de todos aunque no exista un piso económico para ellas.

No está clara la justificación de una eventual franja de radio para candidatos presidenciales, porque ellos ya reciben la franja de TV abierta.

4. El subsidio propuesto para la operación de los partidos ofrece pagar un 50% de los gastos comprobados en funcionamiento regular de los partidos que obtengan suficientes votos en las elecciones como para conservar su inscripción de tal (lo cual significa obtener al menos 5% de los votos nacionales, en términos generales).

Considerando que 90% de los votos válidamente emitidos llegan a partidos que sobreviven, que los votos válidamente emitidos en las parlamentarias son 7 millones, que cada uno eleva el límite máximo para el subsidio en 0,003 UF al mes, y que la tasa de utilización de este subsidio es de 70% (recordar que el partido debe financiar el 50% de sus gastos con otras fuentes), y que cada 48 meses hay 6 meses cercanos al período de campaña sin acceso al subsidio, el costo anual estimado es:

$$G2 = 0,90 \times 7 \text{ millones} \times 0,003 \times \$16.783,6 \times 0,70 \times 12 \times (48-6/48) = 2.331 \text{ millones/año.}$$

5. El subsidio propuesto para las cuotas de militantes que voten en las internas y tengan las elecciones al día es de 30% con tope de 3 UTM al año (\$90.000/año), y a una tasa de 10% para las cuotas mayores (más baja para desincentivar la evasión). Según los antecedentes expuestos en Donoso *et al.* (2000), no es realista pensar en cuotas mayores de \$2,5 mil al mes.

Considerando 100 mil militantes con esas características en todos los partidos, que en promedio donen \$2.500 al mes (\$30.000 al año), resulta un costo fiscal de:

$$G3 = 0,30 \times 100.000 \times \$30.000 = \$ 900 \text{ millones/año.}$$

6. El subsidio propuesto a la capacitación política consiste en subsidiar hasta 30 actividades anuales por cada distrito electoral en 2 UTM por hora, con al menos 12 horas por actividad cada una. Considerando 60 distritos, 30 actividades anuales por distrito, y un subsidio de $2 \times 12 = 24$ UTM (\$720.000) por actividad, resulta un costo fiscal total de:

$$G4 = 60 \times 30 \times \$720.000 = 1.296 \text{ millones/año.}$$

7. El saldo neto de todos estos costos fiscales alcanza a:
\$ 4.842 millones al año (0,015% del PIB o 6,6 millones de dólares al año).

Un diezmilésimo del PIB (0,01%) equivale a 4.500 millones de pesos al año.

Esta cifra no toma en cuenta el mayor presupuesto que debe otorgarse al Servicio Electoral. No se entrega aquí una cifra por falta de antecedentes. ☐