

REVISION DE LIBROS

Nuevos modelos para la medición de actitudes

A. J. Rojas y C. Pérez Meléndez

Promolibro, Valencia, 2001, 214 pp.

No hay mucha literatura en castellano sobre la Teoría de la Respuesta al Ítem (TRI), pero aún hay menos sobre los modelos politómicos desarrollados en el marco de la TRI. En los últimos años parece que están despertando más interés, como muestran, por ejemplo, las tesis doctorales de Espejo (1998), Revuelta (2000) y Abad (2001).

La TRI para ítems dicotómicos se ha ido abriendo camino en nuestras universidades y en la investigación psicométrica. Para la medición de buena parte de los rasgos de personalidad, actitudes... se utilizan ítems con varias categorías (ítems politómicos). La TRI para ítems politómicos permite sacar partido de su solidez y ventajas a la hora de medir estos rasgos. El libro trata de la medición de rasgos mediante ítems en los que las categorías de respuesta son más de dos y están ordenadas; es decir, básicamente, si bien no exclusivamente, de los ítems tipo Likert. Los modelos politómicos para variables nominales no son estudiados en el libro. De los distintos modelos propuestos para el análisis de los ítems con categorías ordenadas, los que van a recibir atención son el modelo de crédito parcial (MCP) y el de escalas de clasificación (MEC). Ambos pertenecen a la familia de modelos de Rasch. Una revisión de estos y otros modelos politómicos puede encontrarse en van der Linden y Hambleton (1997).

Los autores, Antonio Rojas y Cristino Pérez, son profesores de Metodología de las Ciencias del Comportamiento, de las universidades de Almería y Granada, respectivamente. Conocen bien a nuestros estudiantes y el libro va dirigido a ellos. Tiene una orientación más aplicada que teórica. En sus propias palabras: «Este trabajo muestra una mayor preocupación por los aspectos psicométricos que por los matemáticos-estadísticos, queriendo esto decir que nos centraremos más en el significado psicológico-psicométrico que en el desarrollo de los algoritmos matemáticos que los sustentan» (p. 15).

El libro consta de ocho capítulos y una introducción. El primer capítulo, además de exponer algunos conceptos básicos (constructo, variable...), trata de los requisitos métricos de los modelos de la TRI. En particular, se expone la relación entre estos modelos y dos teorías de la medición: la representacional y la medida conjunta. El capítulo no resulta fácil de leer y quizás se debiera advertir al lector de la conveniencia de que lo lea al final, pues su contenido no es necesario para entender los capítulos que siguen.

El segundo capítulo resulta más familiar a todos los interesados en el estudio de las actitudes, pues trata de los distintos formatos de ítems y sus correspondientes normas de cuantificación. Se comentan las principales características de los ítems tipo Likert. Se introduce el importante concepto de «paso». Un ítem con

«n» categorías ordenadas, supone «n-1» pasos (de la primera a la segunda, de la segunda a la tercera...). El lector puede echar en falta alguna reflexión o referencia bibliográfica sobre el número recomendable de categorías, la conveniencia o no de la categoría central, la importancia de las etiquetas asignadas a las categorías... Si tiene paciencia, encontrará respuesta a su inquietud en el capítulo 4.

El capítulo tercero lleva por título «modelos basados en la TRI y la familia de modelos de Rasch». Comienza introduciendo el concepto de «invarianza» o «medida objetiva», que es una de las principales ventajas de la TRI: El cumplimiento del principio de invarianza nos asegura que los parámetros de los ítems no dependen de las personas a las que se han aplicado, y que sus niveles de rasgo tampoco dependen de los ítems que han resuelto. A continuación, se hace referencia a las ventajas adicionales (por ejemplo, en relación a las propiedades de los estimadores) que tienen los modelos que pertenecen a la familia de modelos de Rasch, como les ocurre al MCP y MEC. Resulta ilustrativa la tabla en la que se muestran los principales modelos clasificados por el tipo de ítem y por su pertenencia o no a la familia de Rasch, así como la discusión sobre las características de unos y otros. El capítulo acaba con los siguientes cinco contenidos: a) los supuestos de los modelos de Rasch, b) el modelo de Rasch para ítems dicotómicos, c) los conceptos de curva característica de los pasos del ítem y curva característica de las categorías de respuestas, d) la expresión general del modelo politómico de Rasch, y e) un primer acercamiento al significado de los parámetros de los ítems en dicho modelo.

El capítulo cuarto se ocupa de la exposición de los modelos de crédito parcial (MCP) y de escalas de clasificación (MEC). Del primero se comentan dos ejemplos que muestran bien las consecuencias de que los parámetros de los pasos del ítem estén o no ordenados. A lo largo del texto se exponen muchas gráficas, que facilitan su comprensión. Este es el caso de las mostradas en este tema. A continuación, se introduce el MEC como un caso particular del MCP, que es especialmente indicado para tests en los que el formato de respuesta es común para todos los ítems. De hecho, mientras en el MCP se estima en cada ítem un parámetro para cada paso, en el MEC los parámetros de umbral (diferencias entre los parámetros del cada paso y el parámetro del ítem) han de ser los mismos en todos los ítems del test, por lo que en este modelo se ha de estimar un parámetro por ítem y otro por cada paso. Por cierto, no queda del todo claro si en cada ítem la suma de los parámetros de umbral ha de ser cero, o si, como se indica en el siguiente capítulo, en la estimación se fija arbitrariamente el valor de uno de ellos. El capítulo termina con la descripción de las regiones de respuesta más probable y la puntuación esperada. A partir de los parámetros de los pasos del ítem (o de los parámetros de umbral) se puede obtener fácilmente cual es la categoría de respuesta más probable y la categoría de respuesta esperada para cada nivel de rasgo. Esta información resulta muy útil para conocer si las categorías establecidas en los ítems están funcionando de manera correcta o no.

Los capítulos primero y quinto han sido escritos por los profesores González Gómez y Padilla García. El quinto se ocupa de la estimación de parámetros. El lector queda algo impresionado, y teme lo peor, cuando lee que en la TRI «la estimación de parámetros es un paso ineludible pero lamentablemente escabroso» (p. 93). No es fácil explicar con palabras, casi sin fórmulas, los diferentes procedimientos de estimación, y no es especialmente clara la exposición de los procedimientos de máxima-verosimilitud conjunta, condicionada y marginal. Sin embargo, se explica bien el sentido de la estimación máximo-verosímil. Tampoco se describe claramente la relación que guarda la estimación «PROX» con los procedimientos de estimación ordinarios, comentados en las páginas precedentes. Se entiende mejor la explicación del error de estimación, los procedimientos de bondad de ajuste y el concepto de función de información, especialmente este último.

El siguiente capítulo, «Aplicación de los modelos», describe detalladamente todo el proceso conducente a determinar los ítems que han de componer la prueba, y es probablemente el capítulo más útil para el lector interesado en aplicar los modelos politómicos a la construcción de pruebas psicológicas. Se explica, con datos reales, en varios pasos, cómo eliminar los ítems y sujetos que no se ajustan bien al modelo. Es bien conocido que varios estadísticos de ajuste dependen del tamaño muestral. Por ello, los autores aplican un procedimiento consistente en decidir los ítems a retener tras la comparación simultánea de varias submuestras, de distinto tamaño, extraídas de la muestra total, y no solo, como es habitual, a partir de los resultados de la muestra total. Es también reseñable el extenso y cuidadoso estudio sobre el ajuste del test final, en el que se analizan varios índices de ajuste de manera simultánea. Se describe con especial cuidado y acierto a qué se debe el desajuste de ítems y personas. Calibrado el banco, el siguiente paso es obtener información sobre los ítems y el test en su conjunto, a partir de sus parámetros, y medir a las personas. Los autores muestran detenidamente la información que ofrece el programa BIGSTEPS tras la calibración de cada ítem y el significado de cada elemento de la salida. Para la interpretación de las puntuaciones de las personas, BIGSTEPS ofrece la tabla de equivalencias entre las puntuaciones directas en el test y el nivel de rasgo que corresponde a cada puntuación directa. También informa del correspondiente error de estimación. Termina el capítulo mostrando dos gráficas: la que relaciona las puntuaciones directas con las puntuaciones en el rasgo y la que ubica en el continuo del rasgo conjuntamente a las personas y a los ítems.

En el capítulo 7 se describen varias aplicaciones de los modelos TRI politómicos, como la construcción de tests, la creación de bancos de ítems, los tests a medida y el estudio del sesgo. Los comentarios se hacen en relación a los modelos MCP y MEC, pero lo dicho vale para otros modelos politómicos e incluso dicotómicos. Se explica con detalle la selección de ítems cuando se pretende satisfacer una función de información del test objetivo y se comenta brevemente las complicaciones que pueden derivarse de tal proceder; por ejemplo, las relativas al incumplimiento del plan de especificaciones, con su consiguiente riesgo para la validez del test.

En un breve capítulo final se citan los programas de ordenador disponibles y se comenta con algo más de detalle el programa BIGSTEPS. No se entra a fondo en las ventajas y desventajas de cada uno.

Coincido plenamente con José Muñiz, en sus palabras del Prólogo, en dar la bienvenida al libro sin reservas. El libro tienes varios méritos a destacar: La presencia de este libro en nuestras las

librerías va a permitir el acercamiento a los modelos politómicos de la TRI de personas que muy probablemente nunca a lo harían a los escritos originales en inglés y de tratamiento más teórico. Otro mérito del libro es que resulta realmente motivador e incita al uso de los modelos estudiados, pues muestra muy bien lo útil que son para conocer cómo está funcionando un ítem, sus categorías de respuesta, tal o cual sujeto, para interpretar el nivel de desajuste... Otro aspecto positivo es que está en general bien escrito, aunque tiene algunas erratas. También hay que destacar el esfuerzo que los autores han hecho por hacer comprensibles al lector los contenidos tratados, que no son especialmente fáciles de explicar. Entre los aspectos menos positivos, ya hemos señalado algunos asuntos que no parecen especialmente bien explicados. Se echa en falta una adecuada explicación de los «logits». Se observa también alguna inconsistencia en notación entre el capítulo quinto y los demás.

La aparición, entre otros, de los libros de Muñiz (1990), Santisteban (1990) y López Pina (1995) impulsó el estudio de la TRI en modelos dicotómicos. El libro de Rojas y Pérez debe tener un papel similar en lo relativo a los modelos politómicos.

Referencias

- Abad, F.J. (2001). *Algunas soluciones para la estimación del modelo de elección múltiple de Thissen y Steinberg*. Tesis doctoral. Universidad Autónoma de Madrid.
- Espejo, B. (1998). *Escalas de respuesta y sistemas de puntuación: un análisis mediante modelos politómicos de la Teoría de la Respuesta al Ítem*. Tesis doctoral. Universidad de Valencia.
- López-Pina, J.A. (1995). *Teoría del Respuesta al Ítem: Fundamentos*. Barcelona: PPU.
- Muñiz, J. (1990). *Teoría de Respuesta a los Ítems*. Madrid: Pirámide.
- Reuelta, J. (2000). *A psychometric model for multiple choice items*. Tesis doctoral. Universidad Autónoma de Madrid.
- Santisteban, C. (1990). *Psicometría*. Madrid: Norma.
- Van der Linden, W. J. y Hambleton, R.K. (1997). *Handbook of modern Item Response Theory*. New York: Springer-Verlag.

Revisado por:

???

Universidad de

Métodos, diseños y técnicas de investigación psicológica

María José Navas Ara

Universidad Nacional de Educación a Distancia, Madrid, 2001. ISBN: 84-362-4496-6

Este libro ofrece una introducción general a la investigación en Psicología y Educación, abarcando los fundamentos de la investigación científica, los aspectos psicométricos esenciales relativos a la construcción y valoración de pruebas y las familias de diseños de investigación más importantes. Está diseñado para cubrir los contenidos de la asignatura del mismo título de la licenciatura de

Psicopedagogía de la UNED y consecuentemente, ha sido estructurado en su contenido de un modo adecuado a ese propósito y con los recursos didácticos propios para la enseñanza no presencial, recursos didácticos que, dicho sea de paso, nunca están de más en las «universidades con aulas». El libro constituye una útil y bien seleccionada introducción a los aspectos psicométricos y de diseño de la investigación psicológica, con la atención centrada en el área psicoeducativa.

El libro está estructurado en tres partes: 1) Fundamentos de la Metodología Científica, 2) Construcción, Evaluación y Aplicación de Instrumentos de Medida y 3) Métodos de Investigación. En la primera parte, la más breve, se introduce la lógica de la investigación científica y la medición de lo psicológico. En la segunda, la más extensa, se presenta la construcción y aplicación de los instrumentos de medición psicológica, la asignación e interpretación de puntuaciones, las teorías de los tests, el análisis de los ítems, la fiabilidad y la validez. La tercera y última parte describe la metodología experimental, cuasi experimental, de encuestas, observacional y cualitativa. Las tres partes aparecen integradas entre sí, dentro de una secuencia transparente en la que se hace consciente al lector del hilo conductor y de las frecuentes conexiones entre los diversos ámbitos metodológicos abordados.

La primera parte, dedicada a los fundamentos de la metodología científica, consta de dos capítulos, firmados ambos por la coordinadora del libro. En ellos se caracteriza el método científico y la medición psicológica. Estos capítulos abordan lo que puede considerarse estrictamente esencial de la lógica de la investigación científica y de la medición de lo psicológico, con claridad positiva, de un modo sencillo y próximo, pero sin simplificaciones innecesarias. En ellos el lector encontrará una primera aproximación, esclarecedora y oportuna, a estas cuestiones esenciales. Se recorren los principales tópicos tradicionales, como la teoría de las escalas de medida de Stevens o la clasificación de los instrumentos de medida, pero también se analizan las dificultades de la medición y se insinúan las complejidades que hacen interesante esta temática. La ausencia, tan usual en los actuales planes de estudio, de un tratamiento más explícito y profundo de los fundamentos del conocimiento científico hace todavía más acertado que este tipo de cuestiones sean tratadas en un manual orientado a ofrecer los fundamentos de una formación metodológica.

La segunda parte se ocupa de la construcción, evaluación y aplicación de instrumentos de medida. El capítulo 3 se ocupa de la construcción de tests. Descompone el proceso en cuatro fases: objetivo del test, especificación del test, confección de los ítems y confección de la prueba. El capítulo analiza puntualmente cada una de estas fases con una orientación eminentemente aplicada y ofreciendo suficientes detalles para hacer el material útil en la futura práctica profesional de un estudiante de psicopedagogía.

El capítulo 4 mantiene la transparencia y funcionalidad de los precedentes. Se ocupa de la puntuación de las respuestas a los ítems, de la combinación de esas puntuaciones y de su interpretación, tanto en su modalidad normativa como criterial.

El capítulo 5 se ocupa sucintamente de las teorías de los tests. Se plantea la cuestión del error de medida, se introduce el modelo de la teoría clásica de los tests y algunas de las nociones básicas de la teoría de la respuesta al ítem. El valor didáctico de algunos pasajes, como el que relaciona las tallas de ropa con los supuestos de los métodos de estimación mediante intervalos de confianza, no tiene desperdicio. El capítulo es conciso y limpio, cuenta lo esencial y lo hace con naturalidad, de un modo muy asequible.

El capítulo 6 se ocupa de la cuestión de la fiabilidad. Se presenta la concepción de la fiabilidad de la teoría clásica, en sus modalidades de estabilidad temporal (donde se opta por incluir el método de las formas paralelas junto al test-retest) y de consistencia interna (abarcando los métodos basados en la división del test en dos mitades y el coeficiente alfa), así como dos de los factores que influyen en el coeficiente de fiabilidad, la longitud del test y la variabilidad de la muestra de sujetos. El capítulo también presenta muy brevemente la concepción de la fiabilidad como información propia de la teoría de la respuesta al ítem, y con algo más de detalle la fiabilidad interjueces y la fiabilidad como consistencia en las clasificaciones obtenidas a partir del test, incluyendo en esta última la problemática específica de la interpretación criterial, la precisión del punto de corte y la consistencia de las clasificaciones.

El capítulo 7 se dedica a la validez, después de una introducción general, los contenidos se estructuran en torno a cuatro puntos: métodos relacionados con la estructura interna del test, métodos relacionados con el contenido del test y métodos basados en la relación con otras variables, donde se considera la predicción de un criterio tanto continuo como discreto. Un apartado final se dedica a la cuestión de la generalización de la validez.

Una de las virtudes del capítulo es no reducir la relación test-criterio al caso lineal para variables continuas, explorando también las técnicas adecuadas para criterios discretos con tests continuos (p.e. Análisis discriminante o regresión logística) o discretos (p.e. modelos LOGIT, PROBIT). El problema de la restricción del rango está tratado con una afortunada aplicación de los supuestos que resulta tan didáctica como reveladora del significado del método.

El capítulo 7 es el último firmado por la profesora María José Navas, coordinadora del volumen. Estos siete capítulos tienen en común la sencillez y elegancia de estilo que hace al lector fácil lo difícil y que sólo se adquiere con un conocimiento sereno y bien sopesado. A partir del octavo los capítulos han sido encargados a diferentes especialistas, con las inevitables diferencias de estilo, aunque el carácter introductorio y didáctico del texto puede considerarse común a todos ellos y el esfuerzo de equilibrio temático e integración de contenidos permanece patente en toda la obra.

El capítulo 8, firmado por Juan Carlos Suárez Falcón se ocupa del análisis de ítems, aborda directamente la dificultad del ítem, la discriminación del ítem, el índice de fiabilidad del ítem, el índice de validez del ítem, la relación de los parámetros del ítem con los del test, el análisis de los distractores y el análisis del funcionamiento diferencial del ítem. Se trata de un tratamiento convencional de esta temática donde se presentan sucintamente los tópicos tradicionales para un nivel introductorio.

La tercera y última parte dedicada a métodos de investigación se abre con el capítulo 9, dedicado a la metodología experimental y cuasi experimental. El capítulo, firmado por Ángel M. Fidalgo está bien estructurado en tres partes: conceptos básicos (v.g. hipótesis experimental, causalidad, tipos de variables, optimización, control, modelización, fases de investigación y clases de validez), metodología experimental y metodología cuasi experimental. En el apartado de metodología experimental se abordan los diseños completamente aleatorizados, los diseños de bloque homogéneos aleatorizados, los diseños de medidas repetidas y los de caso único o $n=1$ (quizás mejor que de «sujeto» único, pues como bien se muestra en el texto, el caso único no siempre es propiamente un sujeto). En el apartado de metodología cuasi experimental se abordan los diseños de grupo control no equivalente, los diseños de discontinuidad en la regresión y los diseños de series temporales

interrumpidas. El capítulo, que se inicia con una inusual, aunque atractiva, invocación a los dioses, termina en otra invocación, también inusual pero bien terrenal, al papel insustituible del experimentador. Se trata de un capítulo lúcido, dirigido al entendimiento más que a la memoria. Escrito en un estilo característico y particular de su autor, inusual para un texto científico en estas décadas descreídas y asépticas, pero sabroso. Que hace más cómplice al lector y más placentera la lectura, lo que es mucho, hablando de un texto de metodología. En todo caso, para los de apetitos convencionales que no gusten de estas pequeñas vanidades en el estilo, convendrán conmigo en que estas son difíciles de reprochar cuando en nada perjudican al rigor y más allá, traslucen como aquí, esa comprensión profunda de lo que se explica que es tan inusual como gratificante.

El capítulo 10, firmado por Guillermo Gil Escudero y Rosario Martínez Arias se dedica a la metodología de encuestas. Se introduce el concepto de encuesta y se presenta una clasificación de los principales tipos de encuestas. En los puntos siguientes del capítulo se presentan con cierto detalle las fases de la realización de una encuesta, se abordan brevemente los procedimientos de recogida de información y se presentan las principales técnicas de muestreo, tratando siempre de aproximar aquellos aspectos más propios del ámbito educativo de acuerdo con el marco de referencia general del texto. Entre los métodos de muestreo se presenta el muestreo aleatorio simple, el muestreo aleatorio estratificado, el muestreo aleatorio por conglomerados y los métodos no probabilísticos. El capítulo brinda una visión equilibrada y útil de esta temática.

El capítulo 11, dedicado a la metodología observacional, es obra de Encarnación Sarriá Sánchez y Ángeles Brioso Díez. El capítulo presenta una caracterización del método observacional, presta una adecuada atención a la cuestión de la categorización, explica la cuestión de las reglas de muestreo y las reglas de registro, la métrica de la observación, el control de la calidad de los datos y los diseños y análisis de datos. Se trata de una presentación clara, bien articulada, con el énfasis bien repartido en los aspectos importantes.

El capítulo 12 es obra de Concepción Gabriel Esteban y se dedica a la metodología cualitativa. El capítulo parte de uno de los inolvidables pasajes de Saint Exupéry para hacer patente de un modo difícilmente refutable como el enfoque cualitativo puede ser considerado como una mirada complementaria a la que ofrece el enfoque cuantitativo al que se han dedicado los once capítulos precedentes. El capítulo se centra en tres técnicas cualitativas bien reputadas en otras áreas de investigación psicosocial (v.g. en el marketing y la investigación de mercados): la entrevista en profundidad, el grupo de discusión y el grupo triangular. La presentación resultará útil y estimulará una aproximación a los métodos cualitativos más generosa, más abierta a sus aportaciones de las que puede beneficiarse también la aplicación de la metodología cuantitativa y la interpretación de sus resultados.

En este libro deben destacarse cualidades importantes para su cometido docente tanto en su estructura y contenidos como en su estilo y su formato.

Ante todo, se trata de una obra integradora en su propósito y contenidos y bien integrada en su ejecución y metodología. Integrada porque, a pesar de tratarse de una obra coordinada con la participación de varios autores, se aprecia una estructura común, anunciada explícitamente desde el capítulo 3, que hilvana bien los capítulos y les confiere un sentido compartido. Integrada porque el énfasis docente, aunque riguroso, dota a toda la obra de una textura bien armonizada que permite una lectura no sólo útil, también ame-

na y bien enlazada a las necesidades formativas del contexto psicoeducativo. Integrada también en los conocimientos. Una muestra clara de ello es la conexión a través de capítulos de las explicaciones sobre la aplicación de estadísticos. Así por ejemplo el coeficiente k de Cohen es una medida de acuerdo entre clasificaciones que se explica oportunamente para el caso de juicios de jueces, pero cuya aplicación se retoma en varios capítulos, p.e. en el caso de aplicaciones distintas de la misma prueba a un mismo grupo de sujetos o en el caso de la clasificación de sujetos atendiendo a sus puntuaciones en un test y criterio dicotómicos o dicotomizados.

La obra resulta integradora porque frente a la no poco común deshilazón con que los fundamentos de la investigación, la psicometría, el diseño y hasta la estadística, son tratados en los planes de estudio y frente a la usual e inevitable especialización de los materiales docentes en alguno de estos campos, aquí se tienden algunos de los nexos de unión que han de permitir a los estudiantes adquirir el significado conjunto de estas vertientes de la metodología. El estudiante gana así una comprensión significativa de la «gestalt» metodológica. Los fundamentos dan sentido a lo psicométrico y al diseño. Lo psicométrico adquiere su lugar en el diseño y éste muestra su enraizamiento en lo psicométrico.

La obra también resulta integradora por recoger conjuntamente aproximaciones metodológicas que suelen tratarse separadamente. Así debe destacarse el que se preste atención a la metodología observacional, a la metodología de encuestas y, especialmente, a la metodología cualitativa. Incluso para el lector de metodología más convencional, resultará difícil sustraerse a los motivos aducidos para incorporar a los procedimientos de trabajo esta «otra mirada» cualitativa. Tal como se propone, complementaria, generadora de hipótesis, reveladora de significado.

Todos los conceptos básicos necesarios para seguir el texto, incluso los de naturaleza estadística más elemental como la distribución normal o el modelo de regresión, se explican brevemente y de un modo sencillo. De esta forma no se presupone al estudiante una formación estadística particular, cualidad inapreciable si se espera un aprovechamiento docente que supere la habitualmente pronunciada variabilidad del grado de formación inicial de los estudiantes.

Subrayando el enfoque integrador del libro se introducen en los capítulos de psicometría algunas nociones principales de la teoría de la respuesta al ítem presentadas de un modo escueto, anotando los conceptos básicos. Quizás aquí, la dificultad intrínseca de este cuerpo teórico para un lector como el que presupone el texto, requerirá la necesidad de explicaciones adicionales en el quehacer docente. Al lector se le ilustra sobre la existencia de esta aproximación, lo que quizá estimule el impulso hacia un ahondamiento de mayor calado que el lector habrá de satisfacer con la ayuda de recursos didácticos adicionales.

El planteamiento del texto, la coordinación del mismo y el estilo de la profesora Navas es irreprochable. La sencillez y espontaneidad con que el lector es llevado a través de los temas, sin obstáculo, de un modo diáfano, con buen sentido didáctico, hace de este manual una adquisición inevitable para todas las bibliotecas de Psicología y una herramienta muy recomendable para la formación en los diversos ámbitos de se ocupa, incluso en aquellos planes de estudio en que estas partes son abordadas en materias separadas de un modo más especializado.

Revisado por:
Josep L. Melià
 Universitat de València

IQ and the wealth of nations (CI y la riqueza de las naciones)

Richard Lynn y Tatu Vanhanen

London, Praeger

Richard Lynn es profesor emérito de Psicología en la Universidad del Ulster (Irlanda del Norte). Tatu Vanhanen es profesor emérito de ciencia política en la Universidad de Tampere (Finlandia). Lynn y Vanhanen contrastan en su libro una teoría nunca propuesta con anterioridad: *la inteligencia de las poblaciones es una causa de las diferencias de desarrollo económico que separan a las naciones*.

En el primer capítulo se preguntan por qué algunos países son pobres y otros son ricos, presentando teorías tales como las climáticas y geográficas, las de la modernización, las psicológicas, las culturales, las basadas en el sistema mundial, las neoclásicas o las multi-causales. Los autores señalan que estas teorías explican solo parcialmente las disparidades nacionales, puesto que se olvidan del factor clave: la inteligencia de las poblaciones.

El segundo capítulo introduce el concepto de inteligencia. Hacen notar que el supuesto de que los ciudadanos de las diferentes naciones poseen el mismo nivel de inteligencia «es seriamente incorrecto» (p. 23). Las evidencias disponibles indican que los japoneses, chinos y otros orientales poseen un CI medio de 105. Los europeos de Europa, Canadá, Australia, Nueva Zelanda, Latinoamérica y Sudáfrica, presentan un CI medio de 100. Las personas del sur y sudoeste de Asia –desde Turquía a la India– presentan un CI entre 78 y 90. Las personas de Latinoamérica poseen un CI medio de 96 en Argentina y Uruguay, pero de 80 en países como Guatemala y Ecuador. Finalmente, las naciones sub-saharianas presentan un CI entre 65 y 75. La teoría de Lynn y Vanhanen propone que ya que existe una serie de asociaciones demostradas entre CI y varios fenómenos sociales –incluyendo las diferencias de ingresos– dentro de algunos países europeos o de los Estados Unidos, las diferencias nacionales de inteligencia pueden estar también relacionadas con sus diferencias de riqueza. El tercer capítulo presenta evidencias concretas sobre la relación entre CI e ingresos: «la inteligencia determina los ingresos, ya que las personas más inteligentes aprenden más rápido, resuelven problemas con mayor eficacia, pueden ser entrenadas para adquirir habilidades más complejas, y trabajan de un modo más eficiente y productivo» (p. 35). El cuarto capítulo es una continuación del tercero, revisándose evidencias conocidas sobre la relación del CI con el nivel educativo, el nivel socioeconómico, los logros vitales, la genialidad y la serie de fenómenos sociales explorados en el famoso y controvertido libro de Herrnstein y Murray (*The Bell Curve*). Concluyen que las naciones con mayor nivel de inteligencia se educarán más y poseerán un mayor número de ciudadanos que puedan contribuir al progreso económico de su país. El quinto capítulo revisa relaciones entre inteligencia y desarrollo económico dentro de una serie de países. Los ejemplos corresponden a Estados Unidos, Gran Bretaña y España. Las regiones de estos países con mayor nivel de inteligencia, presentan también mayor desarrollo económico. Lynn y Vanhanen consideran que las regiones están compuestas por *agregados de individuos*. Las naciones también.

En el capítulo seis comienza el análisis propio de los autores. Sostienen que su teoría es contrastable. Basta con conocer el CI medio de las naciones a estudiar y su éxito económico. Existen medidas directas del CI en 81 países. El CI de los británicos se usa como referencia (CI = 100) calculando el CI de los restantes países a partir de ese valor. La medida de inteligencia usada es el test de Raven en la mayor parte de los casos. El CI de los 104 países restantes que completan el globo terráqueo se estima usando el CI de los países colindantes. En este capítulo se intenta demostrar que el cálculo del nivel nacional de inteligencia resulta fiable y válido. Por ejemplo, la correlación entre el nivel de CI de 38 países y el rendimiento en las pruebas internacionales sobre nivel de conocimientos en ciencia y matemáticas que comparan a esos países anualmente, está por encima de +.86. Por otro lado, los datos sobre desarrollo económico se extraen de varias fuentes (p.e. los informes del banco mundial) aunque los diferentes indicadores usados presentan correlaciones superiores a +.94. En los análisis que están por venir se correlacionan las diferencias económicas con las medidas de inteligencia. Además, se calculan varios análisis de regresión para poder ubicar a los distintos países y estudiar a los que más se desvían de la línea de regresión. Este segundo análisis permite detectar otras causas *sistemáticas* de las diferencias económicas que separan a los países.

Los capítulos siete y ocho presentan los resultados de los análisis para el grupo de 81 países en los que existen medidas directas del CI y para el grupo general de 185 países. Las correlaciones entre inteligencia y riqueza se calculan desde que existen datos, es decir, desde 1820 hasta 1998. Un resultado llamativo es que las diferencias nacionales de inteligencia predicen la persistente diferencia en riqueza que ha separado a las naciones desde hace casi dos siglos. Además, la capacidad predictiva del CI ha aumentado: 29% de la varianza en 1820 y 52% de la varianza en 1998. Cuando se analiza individualmente a los países, se observa que la frecuencia de países ricos se corresponde estrechamente con aquellos que poseen un CI por encima de 90. A partir de aquí se preguntan los autores: ¿es necesario un CI de 90 o superior para adoptar efectivamente las nuevas tecnologías que estimulan el desarrollo económico?

Los valores señalados para los 81 países, se replican fielmente para los 185 países. Las diferencias de CI explican un 50% de las diferencias nacionales de riqueza. Ningún otro factor considerado hasta ahora ha logrado explicar tanta varianza. Sin embargo, ese 50% deja margen para la intervención de otros factores. El análisis de regresión permite explorar esta cuestión. Analizan una serie de países en 1900, 1930, 1960 y 1998. En 1900, los casos de China, Italia, Japón, Corea del Sur, Rusia, Taiwán y Tailandia constituyen ejemplos que contradicen su teoría. El CI medio en estos países está por encima de 100, pero en 1900 eran países relativamente pobres. Sin embargo, explican la inconsistencia a través del hecho de que la industrialización no pudo asumirse por la presencia de sistemas políticos autoritarios. El análisis de 1960 también es interesante, ya que los autores se preguntan si las secuelas de la Segunda Guerra Mundial distorsionan la relación entre CI y riqueza. La respuesta es negativa. De hecho, usan el caso de Alemania para apoyar su teoría: «independientemente de la devastación causada por la guerra, Alemania fue capaz de recuperarse en 15 años e ingresar en el grupo de los países más ricos. Esto indica la significación de la inteligencia general de las poblaciones» (p. 132). También discuten el caso de España. En 1960 presentaba un desarrollo económico menor del esperado según el CI medio nacional,

y, en opinión de los autores, la causa reside en el sistema autoritario establecido tras nuestra guerra civil. Los análisis de 1998 demuestran que en los 100 años previos, únicamente algunos países han logrado cambiar su posición relativa. El capítulo ocho termina con un análisis de países con mayor desarrollo económico que el esperado por su CI, con menor desarrollo y con el desarrollo predecible a partir del nivel nacional de inteligencia (en 1998, España presenta un desarrollo ligeramente superior al esperable según su CI nacional de 97). Los que presentan un alto desarrollo se caracterizan por poseer un CI medio de 93 o superior, una economía de mercado y un sistema democrático. En cualquier caso, en 1998 la correlación entre CI y riqueza es +.71, mientras que la correlación múltiple del CI, la libertad económica y el índice de democratización con la riqueza es +.79. Por tanto, la libertad económica y la democracia contribuyen con 10 puntos porcentuales a la predicción de las diferencias nacionales de riqueza (frente a los 50 puntos porcentuales con los que contribuye el CI).

El capítulo nueve revisa las teorías enumeradas en el primer capítulo teniendo en cuenta la capacidad explicativa de las diferencias nacionales de inteligencia. Hay seis modos de sacar partido a un nivel nacional de inteligencia situado alrededor de la media de 100:

1. Ya que la inteligencia predice los logros educativos, los niños de las naciones de alta inteligencia presentan un buen rendimiento escolar y adquieren una alta educación, lo que propicia el capital humano necesario para el desarrollo económico.
2. Las naciones de alto CI poseen una elite científica con capacidad suficiente para producir nuevos productos económicamente valiosos.
3. Las naciones de alto CI producen servicios y bienes valorados en el mercado internacional.
4. Las naciones de alto CI tienen un gran número de personas de moderada y alta inteligencia capaces de realizar las funciones de gestión y el tipo de trabajo de alta calidad de los que dependen de una economía saludable.
5. Las naciones de alto CI no tienen un gran número de personas de baja cualificación incapaces de realizar los trabajos necesarios dentro de una economía saludable.
6. Las naciones de bajo CI poseen industrias vinculadas a la agricultura y la minería, poco demandadas en los mercados internacionales.

Según Lynn y Vanhanen, las teorías geográficas y climáticas encajan de modo razonable con las diferencias de riqueza nacional porque la geografía y el clima influyeron durante la evolución de la humanidad sobre las capacidades intelectuales. Las condiciones climáticas se relacionan con el CI medio de las regiones del planeta y afectan al desarrollo económico a través de las diferencias de inteligencia que han surgido durante la evolución. Por otro lado, países asiáticos como Taiwán o Corea del Sur se han comportado en la última parte del Siglo XX como predicen las teorías de la modernización y la convergencia, porque presentan un CI por encima de la media. Sin embargo, los países sub-saharianos o sudamericanos no lo han hecho, debido, en parte, a su bajo CI.

El último capítulo (El Futuro de la Riqueza de las Naciones) extrae consecuencias de la teoría de los autores respecto a cómo mejorar el desarrollo económico de las naciones pobres. Primero de todo, debería incrementarse el nivel medio de inteligencia de esas naciones. El método más efectivo que proponen pasa por me-

jorar la calidad nutricional de las mujeres embarazadas y de los niños. Además, deberían mejorarse sustancialmente los servicios sanitarios. Mejorar la educación no servirá para incrementar la inteligencia de las naciones, pero ayudará a que sus ciudadanos adquieran ciertas habilidades que son económicamente relevantes. También se debería intentar ayudar a las naciones pobres a que mejoren el uso de anticonceptivos. Con todo, sostienen los autores que sería relativamente más fácil que mejorar la inteligencia, cambiar a un sistema político de libre mercado y a sistemas democráticos, factores que también contribuyen claramente al desarrollo económico.

Las implicaciones más importantes del estudio de Lynn y Vanhanen son:

1. El mundo necesita un nuevo código moral internacional basado en el reconocimiento de las diferencias nacionales de inteligencia y las consiguientes desigualdades económicas. Las poblaciones de los países ricos deben aceptar que tienen una obligación ética de ayudar financieramente, de un modo indefinido, a la gente de los países pobres.
2. Hay que continuar con los programas de ayudas a los países pobres, aunque deberían incorporarse programas de mejora de la inteligencia basados en, por ejemplo, las mejoras nutricionales. El reconocimiento de que las diferencias de inteligencia son una causa básica de las diferencias económicas que separan a las naciones hará posible reducir las diferencias de riqueza.

Estamos, en suma, ante una obra novedosa y valiente. Novedosa porque propone una explicación alternativa sobre un fenómeno socialmente tan importante como las diferencias de riqueza nacional. Valiente porque es probable que los «sacerdotes» de la economía y la sociología descarten el planteamiento por *psicologista* (caso cerrado). En promedio, los sociólogos y economistas no suelen aceptar los planteamientos que dan más importancia al individuo que a los sistemas sociales. Sin embargo, a mi entender, la comunidad psicológica podría sentirse satisfecha por la publicación de un libro como este, en el que se intenta ofrecer una explicación psicológica a un fenómeno que nos estaba injustificadamente vedado. Además, debería aplaudirse el esfuerzo de los autores por ofrecer las evidencias necesarias para poder contrastar por cuenta propia el ajuste de su teoría a los datos (hay en el texto principal 41 tablas de datos, así como dos apéndices sobre el cálculo de los CI nacionales y las evidencias sobre desarrollo económico que ocupan más de 70 páginas). La prueba de que algo está cambiando es la creación de un movimiento en la sociología que comienza a mirar con respeto hacia conceptos tan sólidos en psicología como el de inteligencia. La denominada «Sociología de la Inteligencia» explora la capacidad explicativa de las diferencias de inteligencia que separan a los ciudadanos para comprender fenómenos tan relevantes como las diferencias educativas, ocupacionales, económicas o de salud que también separan a los ciudadanos. Lynn & Vanhanen aplican esta lógica al comparar naciones por primera vez en la historia de la ciencia. Está por ver cuál será la reacción a este atractivo, aunque discutible, planteamiento.

Revisado por:
Roberto Colom
 Universidad Autónoma de Madrid