

Importancia del avance en la investigación psicométrica

Nuria Cortada de Kohan
Universidad del Salvador
Universidad de Buenos Aires

Conferencia. XXVIII Congreso Interamericano de Psicología. Santiago de Chile, 29 de Julio al 3 de agosto de 2001.

Resumen

El avance de la investigación en psicometría confirma día a día que las diferencias individuales existen y son medibles. Se señala la necesidad de que el psicólogo que investiga en teoría de los tests tenga un amplio conocimiento de los modelos multivariados además de la estadística descriptiva e inferencial básica debido a la profunda matematización de la psicología. Se analiza detalladamente la necesidad de establecer una estrecha relación entre las teorías cognitivas y la psicometría para que los tests estén al servicio del aprendizaje. Se ha pasado de un sistema selectivo a un sistema adaptativo y como consecuencia es necesario comprender los procesos cognitivos que están involucrados en los tests de aptitudes. Se presentan los aportes de Vygotsky (1978), Hunt (1971,1980), Sternberg (1981,1987), Carroll (1976,1993) y Piaget (1970) y se ejemplifica con algunos tests como el de Kaufman (1985) y el de Woodcock-Muñoz (1996) que enriquecen las posibilidades de evaluación sin despreciar los últimos refinamientos de la técnica psicométrica.

Palabras claves: Psicometría- teorías cognitivas - evaluación del aprendizaje- tests de aptitudes.

Abstract

The advancement of psychometric techniques corroborates every day the existence of individual differences and the possibility of their measurement. It is stressed the need for the psychologist who does research in test theory for a deep knowledge of multivariate designs besides of basic descriptive and inferential statistics. This is due to the great mathematization of Psychology. Today there are three great theories in psychometric methods: 1) The Classical Theory of Tests that is a regression model; 2) the Theory of Generalizability that analyses primarily variance error and uses the concept of facets and the ANOVA and 3) the Item Response Theory that is a probabilistic model used by Birnbaum, Lord, Rasch and some others. But if tests have to benefit learning a connection must be established between cognitive and psychometric theories. Since the onset of testing there have been many changes. From a selective system an adaptative point of view has come out. Now it becomes necessary to understand cognitive processes implicit in the aptitude tests. The cognitive theories in psychology are a dominant point of view in contemporary psychology, not only in the study of intellectual processes but also in social psychology, psychology of personality and developmental psychology. In many instances the structure of mental processes have become the central core of empirical research. The contributions of Vigotsky (1978) Hunt (1971,1980), Sternberg (1981,1987), Carroll (1976, 1993) and Piaget (1970) are analyzed. Some tests as the Kaufman's (1986) and the Woodcock-Muñoz (1996) are considered as assessment tools enriching the measurement. And this is done without laying aside the latest improvements of psychometric methods.

Key words: Psychometric methods - cognitive theories - learning assessment - aptitude tests.

*"Cuando puedas medir aquello de lo que hablas... entonces puedes pensar que sabes algo acerca de ello."
Popular lectures and addresses (1891)
W. Thomson, Lord Kelvin (1824-1907)*

Introducción

La psicología es una ciencia nueva. Si recordamos que el Instituto de Psicología de Wundt en Leipzig fue creado en 1879, que los trabajos de Weber (1795-1878) y Fechner (1801-1887) fueron publicados alrededor de 1860, que Alfredo Binet ideó su escala de edad mental en 1905 y que el primer psicólogo con título de los EEUU. fue James McKeen Cattell que lo obtuvo en 1890, no debe extrañarnos que nuestras carreras de psicología en las universidades latinoamericanas tengan apenas 50 años, aunque en muchos casos precedieron a las de muchos países europeos. Sin embargo, el crecimiento de la psicología ha sido en todo sentido exponencial. Por ejemplo, en el año 1984 la revista *Psychological Abstracts* publicó 32.800 resúmenes y en 1995 se publicaron 47.000, lo que representa una tasa de crecimiento del 42% en diez años. Asimismo, si observamos algunos datos sobre el crecimiento de la cantidad de psicólogos en algunos países como la Argentina, los datos son todavía más impresionantes. En ese país (Alonso, 1996) en 1985 había 20.000 psicólogos y en 1995 la cantidad ascendió a 38.800 lo cual nos da una tasa de aumento del 93%. La cifra significa aproximadamente un psicólogo cada 800 habitantes. Se estima que el 85% de estos eran mujeres. En el año 2001 en la Facultad de Psicología de la Universidad de Buenos Aires están cursando estudios 14.000 alumnos y, además, hay en el país por lo menos otras doce universidades entre públicas y privadas en donde se puede estudiar psicología, aunque todas tienen menor número de alumnos que la de Buenos Aires. Pero una estimación bastante probable es que más del 80% de los egresados se dedica a la psicología clínica.

Y ahora entrando ya en nuestro tema es justo que nos preguntemos: ¿Qué área ha contribuido más para hacer de la psicología una ciencia? Creo no equivocarme al afirmar que ha sido la psicometría que empezó tratando de medir los constructos psicológicos y sigue en la lucha por la matematización de nuestra ciencia para estar a tono con los avances de las demás ciencias y con la complicada tecnología del tercer milenio. En efecto cuando una ciencia tiene por objeto el estudio de procesos dinámicos necesita un aparato formal para analizarlos. La psicología estudia muchos procesos como el aprendizaje, la percepción, la memoria, la atención, los fenómenos cognitivos y emocionales que son procesos que se refieren al cambio o modificación de una variable en función de otras, como por ejemplo la evolución del aprendizaje en función de la práctica, el declive de las huellas de la memoria en función del tiempo, el cambio de niveles de percepción en función de los estímulos, el aumento de las funciones cognitivas en relación al crecimiento, etc. Sin embargo, la psicología clínica y aplicada atrae mucho más a los psicólogos que la psicología experimental y matemática o la teoría de los tests. Tal vez esto se debe en parte a la dificultad de encontrar vocaciones que puedan integrar el interés por las personas con las aptitudes para el pensamiento abstracto, aunque que esto en parte quizás se deba a una deficiencia en la formación de nuestros jóvenes tanto en la escuela secundaria como en la Universidad. Pero en este momento los avances y el potencial en investigación de la psicometría son tan espectaculares que deberían ayudar a cambiar el panorama.

Diferencias individuales

En efecto creo que la teoría de los tests está logrando grandes avances y considero que esto es muy positivo. En un mundo en el que un ejecutivo de Japón, Chicago, Viena o Buenos Aires se viste igual, usa el mismo teléfono celular, cuando es tan fácil comunicarse con Madrid o Estambul, cuando todos los jóvenes sin distinción de sexo o clase social usan *blue-jeans*, donde la moda, las bebidas, los espectáculos, la violencia, los productos son similares en todas partes; en un mundo que globaliza la economía y la uniformidad, la psicometría nos trae alguna esperanza pues cada día nos confirma que todavía el hombre es único y que en nuestra mismidad, como seres biológicos y sociales, cada uno es diferente psicológicamente del otro; que las *diferencias individuales* existen de manera evidente y mensurable.

Estas diferencias ya habían preocupado a los griegos y así Teofrasto en el siglo IV a.C. comentaba: "Hace tiempo que estoy pensando y preocupado por la razón de que todos los griegos tienen un carácter diferente a pesar de vivir en el mismo clima y tener una misma educación". Un poco más cerca nuestro el español Huarte de San Juan que vivió entre 1530 y 1588 en su *Examen de Ingenios para las Ciencias* ya nos habla de las diferencias individuales en las aptitudes humanas y trata de orientar a los jóvenes respetando estas diferencias.

Hace casi cien años nacieron dos grandes programas de investigación específicos en psicología diferencial que partiendo del trasfondo común de la psicología experimental alemana (Fechner 1801-1887) resultan sin embargo diversas en su concepción. Uno fue el programa nacido de la influencia de Galton (1822-1911) en Inglaterra y en Estados Unidos y otro, el de Alfredo Binet en Francia (1857-1911). Galton defendió un esquema científico y tecnológico "hereditarista", partía de un determinismo biológico de la inteligencia que andando el tiempo daría origen a una importante tecnología para poder llegar a la optimización de los recursos humanos. En cambio Binet originaba un pensamiento de tipo ambientalista que daría lugar a un programa para mejorar las capacidades mediante la educación. Sin embargo, ambos puntos de vista propiciaron el desarrollo de los tests para lograr el conocimiento, clasificación y distribución de los talentos humanos utilizando básicamente los mismos instrumentos de medida.

No se puede hacer ahora una descripción histórica del adelanto que esta noción de instrumento de medida ha sido para la psicología. En realidad ha significado el paso de la psicología desde un conocimiento predominantemente filosófico a establecerse como una disciplina cada vez más acorde con los cánones científicos (Koch, 1981).

A pesar de las posiciones que critican los tests por considerar que profundizan poco en el conocimiento de los procesos psicológicos, debemos recordar que nadie ha podido rechazar la afirmación de Meehl (1954) acerca de que la predicción estadística con tests para el diagnóstico era más eficiente que la basada en los juicios clínicos, lo que fue comprobado nuevamente hace pocos años por Dawes (1988). Creo que existe actualmente en algunos psicólogos una especie de pudor al hablar de los tests como de algo ya superado. Sin embargo, la palabra *test* tiene un noble origen pues en el latín medieval el sustantivo *testa-is* del cual proviene, significaba la vasija de barro con la cual los alquimistas ensayaban o probaban la legitimidad del oro (Foulquié, 1967). Por otro lado es evidente que los psicómetros deberían estar más en contacto con otros especialistas de la psicología pues como señala Anastasi la especialización creciente "ha conducido a concentrarse en las técnicas de construcción de tests sin tener bastante en cuenta la investigación psicológica para interpretar los puntajes" (Anastasi, 1967, p. 305).

Matematización de la investigación psicológica

En cuanto a la medición, base de la evaluación psicológica, actualmente un psicólogo no puede investigar en teoría de los tests, si no conoce además de la estadística descriptiva e inferencial básicas una serie de modelos multivariados que hace veinte años eran usados muy raramente. En cuanto a la estadística inferencial, después de los trabajos de Cohen (1988) un psicólogo no se puede conformar hoy en día en señalar que ha hallado una significación con una $p \leq 0.05$ ó 0.01 . Debe informar además sobre el índice del tamaño del efecto de la prueba usado, sobre si se trata de una hipótesis uni o bilateral y del nivel de potencia ($1 - \beta$) que está aceptando para la muestra aleatoria de los n casos usados. La *American Psychological Association* acaba de publicar (Grimm & Yarnold, 1997) un manual sobre técnicas de análisis multivariados en investigación para que sus asociados estén en condiciones de leer los artículos de sus revistas ya que como señalan en el prefacio es "difícil hallar artículos fundamentados empíricamente que no usen algún tipo de análisis multivariado" y entre ellos incluyen análisis factorial exploratorio y confirmatorio, análisis discriminante, path análisis, análisis de correlación y regresión múltiple, escalonamiento multidimensional, elaborados estudios de datos en tablas de contingencia y regresión logística. En efecto a medida que crece el tamaño de los conjuntos de datos la metodología multivariada se vuelve más útil. Las actuales tecnologías facilitan la recolección de grandes cantidades de datos y por eso se necesitan los métodos multivariados para determinar si en realidad esas cantidades de datos contienen información valiosa. El objetivo primario de los análisis multivariados es resumir grandes cantidades de datos en relativamente pocos parámetros. El tema subyacente de muchas técnicas multivariadas es la simplificación. (Johnson, 2000).

Thurstone se mostraba optimista con respecto al progreso de la psicología y comentaba en un artículo publicado en la revista *Science* (Thurstone, 1937, p. 229): "El desarrollo de nuestra ciencia seguirá seguramente el modelo de toda ciencia, haciéndose cada vez más matemática a medida que las ideas se formulen de manera más rigurosa". No sé si entonces podía imaginar cuánto se ha matematizado la psicología de hoy después de los aportes de Torgerson, Stevens, Coombs, Luce, Tversky y varios otros. También gracias al uso masivo que le han conferido las computadoras y los sofisticados *softwares* que facilitan las tareas.

Teorías psicométricas

Pero además de esta matematización de la psicología iniciada y propiciada sobre todo por los grandes teóricos de la psicometría como Spearman, Thurstone, Guilford, Cattell, Cronbach, Lord, Birnbaum, etc. para citar solo algunos, el movimiento de los tests avanza en otro sentido, tal vez menos evidente, pero que a mi entender será muy fructífero y es el de integrar la teoría de los tests con las teorías cognitivas. Hasta ahora la mayoría de los avances en la construcción de los instrumentos se lograba empleando los progresos de la teoría estadística través del uso de ANOVA, el análisis de la dificultad de los ítemes de los índices de dificultad y discriminación, de los ajustes de escalas, de la normalización de los datos, de las correlaciones biserials, de índices más poderoso de confiabilidad y de la obtención de la validez de los constructos merced al análisis factorial, etc. Todo esto ha avanzado y se puede resumir en tres teorías de los tests que son fundamentales y que en parte se complementan.

Por un lado, la más desarrollada es la *Teoría Clásica* iniciada por Spearman que es un modelo de regresión lineal con dos variables cuyo supuesto fundamental es que el puntaje X de una persona en un test es la suma del puntaje verdadero de esta persona más un error: $X = V + e$

Esta es una teoría muy conocida por lo que no insistiremos en ella.

El segundo modelo surgido en los años 60 para complementar el primero se debe a Cronbach y otros y es el llamado de la *Generalizabilidad* que gracias al uso específico del análisis de variancia hace posible analizar a las distintas fuentes de error que se presentan en los puntajes mediante los conceptos *de faceta*. La teoría de la Generalizabilidad expuesta fundamentalmente en la obra de Cronbach, Gleser y otros (1972) siguiendo la línea estadística de Fisher atribuye el concepto de error a múltiples factores de influencia que llaman **facetar**, término introducido por Cronbach para designar cada uno de las características de la situación de medición que pueden modificarse de una ocasión a otra y, por tanto, hacer variar los resultados obtenidos (por ejemplo los ítems del test, las formas de codificar las respuestas, las situaciones de examen, etc.). El diseño de una faceta tendría cuatro fuentes de variación o error. Por ejemplo en un test de rendimiento de matemáticas, la primera fuente sería la diferencia de rendimiento en matemáticas de los estudiantes (objeto de la medición) y lo que se llama variancia verdadera, La segunda fuente sería la dificultad de los ítems, la tercera fuente la interacción de las diferencias de las personas con los ítems y la cuarta fuente los errores aleatorios y fuentes de error no identificados. Estos distintos orígenes del error se discriminan con el ANOVA. Los diseños de una faceta son los más simples, a menudo se incluyen más facetar como fecha de aplicación.

El tercer modelo es el de la Teoría de la Respuesta al Ítem también llamada del rasgo Latente que es un modelo probabilístico y permite conocer la información proporcionada por cada ítem y así crear tests individualizados, es decir a medida. Es un modelo complejo que se ha popularizado como modelo de Rasch(1980) (aunque específicamente el modelo de Rasch es solo de un parámetro la dificultad del ítem), pero existe también el modelo de dos parámetros que tiene en cuenta también la discriminación o pendiente de la curva y el de tres parámetros que tiene en cuenta el factor azar en la respuesta a ítems de alternativas múltiples (Cortada de Kohan, 1999) La diferencia principal entre el modelo clásico y los del rasgo latente (Lord 1980) es que la relación entre el puntaje observado y el rasgo o aptitud en la teoría clásica es de tipo lineal ($X = V + e$) mientras que en los diversos modelos de la TRI las relaciones son funciones de tipo exponencial, principalmente logísticas. Por ejemplo, la expresión para el modelo de tres parámetros es la siguiente:

P

$$i(\theta) = c + eDa(\theta - b) / 1 + eDa(\theta - b).$$

$$i = 1,2,3,\dots,n$$

En donde:

P

$i(\theta)$ = la probabilidad de que un examinado elegido al azar con aptitud θ conteste correctamente el ítem i .

b = parámetro de la dificultad del ítem i

a = parámetro de la *discriminación* del ítem i

n = número de ítems del test

D = 1.7 valor arbitrario introducido para que la función logística sea ajustada a la normal con una exactitud de 0.01.

e= base de los logaritmos neperiano = 2,178

c = parámetro de pseudo azar o sea la probabilidad de contestar bien un ítem de alternativas múltiples por azar.

θ = rasgo latente, aptitud, etc.

Gráficamente esta expresión se representa como la *curva característica del ítem (CCI)*

Los postulados básicos de la TRI son :

- 1) El resultado de un examinado en un ítem puede ser explicado por un conjunto de factores llamados rasgos latentes o aptitudes simbolizados por θ
- 2) La relación entre la respuesta a un ítem y el rasgo latente se describe como una función monótona creciente que es la curva característica del ítem
- 3) En la TRI los parámetros de aptitud y de los ítems son invariantes.

Los supuestos de la TRI son:

- 1) La *unidimensionalidad* del rasgo latente es decir que los ítems de un tests deben medir una sola aptitud o rasgo y
- 2) La *independencia*. Es decir, las respuestas de un examinado a cualquier par de ítems son independientes.

Para estimar los parámetros de los modelos de la TRI se usa el método de *máxima verosimilitud* que es un proceso complejo que se logra con los *software* apropiados como BILOG, BICAL, u otros.

Nosotros hemos sido iniciadores entre algunos otros investigadores (Cortada de Kohan, 1998) del uso de estos modelos en la Argentina. (Galibert, M., Aguerri, E, & Attorresi, H., 2000) He usado el modelo de tres parámetros para un test de aptitud verbal y actualmente estoy creando uno para evaluar conocimientos de matemáticas al terminar los estudios secundarios. En realidad en nuestros países los psicólogos suelen conformarse con hacer los baremos de tests extranjeros adecuando las normas al país. Por nuestro parte, creemos que para muchos objetivos es imprescindible crear instrumentos propios y a esto nos hemos abocado desde hace algunos años.

Integración de las teorías cognitivas y la teoría de los tests

Junto con todo esto y merced al inicio de las teorías cognitivas alrededor de los 50 ahora se está poniendo el acento en la elaboración de los ítems de los tests beneficiándose la psicometría con la investigación sobre la ejecución humana en todos los niveles de competencia. Para que los tests estén al servicio del aprendizaje debe establecerse una relación entre las teorías cognitivas y la psicometría y considerar la evaluación del estudio como una medición de la destreza y aptitudes esenciales para el aprendizaje ulterior. Existen actualmente algunos enfoques diferentes de la evaluación que provienen del cognitivismo y que se centran en el análisis de las tareas, el análisis de los errores y el desarrollo de las competencias (Wittrock & Baker, 1989) Es decir se trata de una evaluación de los procesos.

El punto de vista cognitivo representa un enfoque dominante en la psicología contemporánea no sólo en el estudio de los procesos intelectuales, sino también en la psicología social. En la psicología de la personalidad y en la psicología evolutiva. De

una manera u otra la estructura y procesos de la mente se han convertido en el foco central de los modelos y de las investigaciones empíricas. La psicología cognitiva comprende una amplia gama de enfoques psicológicos que, se dice, juegan un papel fundamental en el comportamiento. La psicología cognitiva reconoce la separación de lo que está *afuera* y su representación interna y considera que el comportamiento es una función del mundo subjetivo que es transformado y representado internamente. Las personas responden al modo como definen las situaciones estimulantes y no a las propiedades objetivas de aquellas funciones estimulantes. Las raíces filosóficas de esta perspectiva surgen de la tradición de Descartes y Kant e implican una *reducción subjetivista* que garantiza la primacía de las estructuras y procesos del sujeto cognoscente y la *reducción individualista* que da prioridad al pensamiento del individuo como es claro en el *cogito* de Descartes . Esta combinación de subjetivismo e individualismo converge para establecer un cuadro de la realidad que logra su coherencia y orden en virtud de la ordenación y universalidad de los elementos de construcción de la mente individual. Es el orden del pensamiento y razonamiento humano el que garantiza el orden y significado del mundo de la realidad (Sampson,1981).

La utilidad de una teoría científica depende de los hechos que explica y de los descubrimientos que inspira. La psicología recibió un gran impacto con las teorías de Newell y Simon(1972) y de G Miller (1967) importadas del ámbito de la computación El problema central que la teoría computacional planteó a la psicología era descubrir la forma de las *representaciones mentales*, es decir, las inscripciones simbólicas utilizadas por la mente y los procesos que las autorizan. La investigación llevada a cabo en el ámbito de la psicología cognitiva ha intentado triangular las representaciones internas de la mente midiendo las informes de los sujetos, los tiempos de reacción, y errores en que incurren cuando recuerdan, resuelven problemas, reconocen objetos y generalizan a partir de la experiencia (Pinker, 2001). Así es como sabemos que nuestra mente contiene representaciones específicas para los conceptos, ese decir las entradas abstractas de las palabras y no para las figuras que cobran las palabras al ser escritas (Pinker,1994). La psicología cognitiva estudia el modo en que las personas representan y procesan mentalmente la información. Sus dominios son la percepción, el aprendizaje, la memoria, la resolución de problemas y la toma de decisiones y aunque los psicólogos cognitivos tienden a subrayar la importancia de la variación de los estímulos y los psicólogos diferenciales ponen el acento en la importancia de la variación individual en ambos casos se están estudiando comportamientos similares. Si los marcos de referencia convergen, la práctica de los tests mentales puede resultar favorecida. Los procedimientos de los tests psicométricos deberían ser complementados con los enfoques de los componentes de los procesos de información, de las representaciones mentales y de las estrategias que la gente usa, analizadas por los cognitivistas.

Según los recientes modelos cognitivistas de la enseñanza, este es el proceso por el cual se conduce a los estudiantes a revisar y mejorar sus convicciones y creencias, es decir el acervo cultural con que comienzan. La enseñanza debe ser mucho más que presentar y desarrollar un tema de estudio. La enseñanza debe incluir evaluar y reconocer el conocimiento y los procesos de pensamiento de los estudiantes y también idear formas de instrucción que contribuyan a este conocimiento. Los procesos de los estudiantes para el aprendizaje de la información tienen papeles decisivos en el desempeño escolar. Las estrategias de aprendizaje de los alumnos tales como leer con comprensión y rapidez, realizar operaciones aritméticas básicas, saber estudiar concentrando la atención, saber repetir lo estudiado en forma escrita y oral y saber recordar fechas, nombres y lugares,etc., influyen mucho sobre el sentido que los estudiantes pueden dar a la información nueva presentada por los profesores por lo que es importante que estos las conozcan y puedan ejercer influencia sobre su

uso. Ahí está probablemente el papel de las evaluaciones diagnósticas de estas estrategias. La más conocida de estas evaluaciones es el test conocido como LASSI (por Learning and Study Strategies Inventory) de Claire Weinstein y otros(1988) que es una escala tipo Likert de 5 puntos, con 90 ítems repartidos en 10 subescalas de grupos estratégicos tales como ansiedad, actitudes, concentración, procesamiento de información, motivación, organización, autoexamen, selección de ideas fundamentales, elementos auxiliares de estudio y estrategias para tomar examen.

Los cognitivistas también le han dado mucha importancia a lo que llaman *metacognición*, esto es, la conciencia y control de los propios procesos de pensamiento. Así los procesos metacognitivos comprenden las formas en que los estudiantes supervisan la comprensión de sus lecturas, planifican como estudiar cada tema y relacionan la nueva información adquirida con sus conocimientos previos. Como parte final, existe lo que se ha dado en llamar procesos de *pensamiento afectivo* tales como la motivación y la ansiedad que contribuyen en gran parte al logro académico. Las nuevas investigaciones demuestran que las *atribuciones* que hacen los estudiantes en cuanto al aprendizaje influyen sobre sus expectativas, sus intereses y persistencia al estudiar y aprender. Por ejemplo, si los estudiantes atribuyen su éxito en la escuela a su esfuerzos, suelen procurar seguir aprendiendo aunque algunas veces fracasen . En cambio si atribuyen el éxito a factores que estén fuera de su control como la suerte o la herencia es difícil que persistan, aún cuando hayan tenido algunos éxitos. Es decir, lo que los alumnos aprenden de los maestros depende del acervo cultural de los alumnos, de sus adecuadas estrategias de aprendizaje, de sus procesos metacognitivos y de sus atribuciones en cuanto al aprendizaje (Wittrock,1990).

Hasta ahora la mayor parte de los avances en la construcción de los tests se lograba empleando los progresos de la teoría estadística a través del análisis de la dificultad de los ítems, de los índices de discriminación de los ajustes a escalas de normalización de los datos, a las correlaciones biserials, a los índices de confiabilidad y a los análisis de ANOVA y de Análisis factorial para la validez de constructo. Pero la teoría cognitiva pretende que sin depreciar estas técnicas, se ponga más el acento en la elaboración de los ítems beneficiándose de la investigación sobre la ejecución humana en los distintos niveles de competencia. Para que los tests estén al servicio del aprendizaje debe establecerse una relación estrecha entre las teorías cognitiva y psicométrica y considerar la evaluación del estudio como una medición de las destrezas y aptitudes esenciales para el aprendizaje ulterior. Cada nivel de conocimiento se debe evaluar de tal forma que permita a los alumnos distinguir como pueden usarlo para conseguir mayor información juzgar evidencias, sopesar alternativas de acción y articular los argumentos. Hay que medir, sobre todo, la capacidad de los alumnos para exponer sus preguntas y expresar sus opiniones (Glaser,1963).

Para educar debemos tener una comprensión clara de lo que los estudiantes saben y de lo que no saben. Además debemos ser capaces de determinar con seguridad si nuestros esfuerzos para instruir han sido eficaces. Así la evaluación de los conocimientos debe proporcionar información que nos ayude a saber si los programas de instrucción son adecuados y si señalan los problemas relacionados con la comprensión y la aplicación de los distintos saberes. Es decir, la evaluación tiene como su más legítima función servir a la enseñanza. Algunos autores (Goetz, Hall & Fresco, 1990)) señalan que los tests y exámenes actuales se centran en la medición gracias a las normas de las diferencias individuales y que las preocupaciones por la validez y la confiabilidad se ha puesto por encima de la necesidad de comprender las naturaleza del aprendizaje y poder así guiar a la instrucción. Además los que construyen los tests y los que los usan para el diagnóstico han sufrido críticas severas

al ser cuestionado el valor de los tests de inteligencia para evaluar el potencial intelectual y la confiabilidad de la clasificación y la asignación a los niveles de instrucción de los alumnos mediante tests. Se consideraba que algunas veces se perpetraban discriminaciones e injusticias, tema ya muy sabido. Actualmente hay menos énfasis en la selección de los individuos para oportunidades educativas y se acentúa más la posición de ayudar a los alumnos a tener éxito en las oportunidades elegidas. Los maestros y las escuelas necesitan tener información sobre los individuos que se oriente hacia la decisión de la instrucción más bien que hacia la predicción. En este sentido, ahora se están evaluando sistemáticamente las influencias de la familia, la escolaridad temprana y las influencias socioculturales sobre los procesos intelectuales que contribuyen a la inteligencia y las aptitudes y no se olvida que los puntajes de los tests están relacionados con las aptitudes adquiridas, las experiencias vitales y las oportunidades educativas. Los puntos de vista han cambiado mucho desde los comienzos de la psicometría. Se ha pasado de un sistema selectivo a un sistema adaptativo, de ayuda y orientado hacia la instrucción. Como consecuencia de ello es urgente comprender los procesos cognitivos que están involucrados en los tests de aptitudes (Glaser, 1981). En este sentido a medida que se van identificando las diferencias individuales en el desarrollo de los procesos cognitivos usados en las tareas de los tests de aptitudes se procede al análisis de las tareas de criterio utilizados para establecer la validez de un test. De esta manera las realizaciones cognitivas que explican la correlación de la aptitud con el criterio se identifican y esto constituye la etapa para construir una teoría de la validez del procesamiento de la información. Es interesante en este aspecto el trabajo de Vygotsky (1978) que establece una distinción en el *nivel de desarrollo actual* de un niño, como el que se puede medir en un test estandarizado y el *nivel de desarrollo potencial* de un niño. Este sería el grado de funcionamiento mental que un niño puede lograr con algún tipo de ayuda y apoyo durante el examen. Ambas medidas son esenciales para el diagnóstico de la aptitud, para el aprendizaje y para los diseños concomitantes de los programas para una instrucción apropiada o de apoyo (*remedial*). Las diferencias entre estas medidas o *zona de desarrollo proximal* como lo llama Vygotsky, se conciben como un indicador del potencial de aprendizaje. Sujetos con el mismo puntaje en un test de aptitud pueden variar en términos de su potencial cognitivo

Debemos mencionar como fundamentales los aportes de Hunt (1980) sobre los estudios de las capacidades verbales en relación con los procesos cognitivos de acceso al léxico, considerado un trabajo pionero en esta línea. Este autor puso de manifiesto que los sujetos con la más alta capacidad verbal en los tests psicométricos, acceden más rápidamente a los códigos en la *memoria de largo plazo (MLP)* para poder establecer los juicios de identidad nominal que los sujetos con poca capacidad verbal (esto se logra experimentalmente en las mediciones del tiempo de reacción, tal como el tiempo de *decisión* para juzgar si dos letras escritas con grafías distintas, por ejemplo A mayúscula y a minúscula son iguales o diferentes. Luego Sternberg (1981, 1987) analizó las capacidades de razonamiento tratando de estudiar la inteligencia según un *análisis componencial*. Para él en vez de ser los factores como para los factorialistas, *el componente* sería el proceso elemental de información que opera sobre las representaciones internas de los objetos o símbolos. El componente puede actuar de tres maneras básicas:

- a) traduciendo un *input* o entrada sensorial en una representación conceptual
- b) transformando una representación conceptual en otra y
- c) traduciendo una representación conceptual en un *output* o salida motriz.

Según la actuación de los componentes y sus distintas clases, Sternberg propone una teoría *triárquica* de la inteligencia que se caracteriza por estimar velocidad de ejecución de los componentes suponiendo que el tiempo que se tarda en resolver un ítem es la suma requerida para ejecutar cada uno de los componentes. La diferencia fundamental entre la teoría de Sternberg y los factorialistas es que lo que organiza el sistema de pensamiento para los factorialistas son las aptitudes y para Sternberg, las tareas. El factor G de Spearman estaría acá representado por lo que Sternberg llama los *metacomponentes* que son procesos de orden elevado que se usan para planificar cómo resolver un problema. Más adelante Carroll (1976,1993) realizó una extensa labor para explicar la naturaleza de los factores en términos de procesos cognitivos.. El trabajo de Carroll es muy complejo y ha usado el modelo de la memoria distributiva de Hunt (1971) que señala la existencia de tres tipos de memoria: a corto, medio y largo plazo, por las que pasa la información, se transforma y se grava o desaparece si no es necesaria. Además Carroll usó la teoría de Neisser (1967) y de Newell(1973) que suponen un proceso ejecutivo que controla todo el flujo de la información estableciendo codificaciones. Finalmente según Carroll, se establece un *programa* o sistema de producción de cualquier tarea cognitiva que comprende operaciones y estrategias. Carroll ha estudiado especialmente lo que denomina *aptitud lingüística* que es la aptitud para el aprendizaje de una segunda lengua aparte de la materna. Carroll ha hecho a través de su teoría un análisis de los componentes de los procesos que conforman los distintos factores que nos permiten describir las diferencias cualitativas de la inteligencia entre las personas.

Piaget (1970), considerado uno de los cognitivistas más relevantes no se interesó por los tests ni por las diferencias individuales (Huteau & Lautrtrey,1997) era un teórico que quería por un lado explicar la génesis del conocimiento científico en la historia de la humanidad (*epistemología genética*) y la génesis del pensamiento lógico en el niño (*psicología genética*). La teoría de Piaget es *estructuralista* dado que según él el hombre no puede comprender el mundo más que asimilándolo a sus estructuras cognitivas y *constructivista* al pensar que estas estructuras, estos instrumentos mentales del conocimiento no son innatos ni sacados el ambiente, sino que se construyen con la coordinación de las acciones. Es decir que al actuar en el mundo para transformarlo, el hombre construye con la coordinación de sus acciones, sistemas de transformación. Estos sistemas son primero realizados materialmente en el período sensoriométrico y luego hay sistemas de acción interiorizados que son *operaciones* del pensamiento. Así surgen en Piaget las grandes etapas del desarrollo del pensamiento lógico: la *etapa sensorio-motriz* (de 0 a 18 meses o 2 años) la etapa *preoperatoria* (de 2 a 7 años aproximadamente), la etapa *operatoria concreta* (de 7 u 8 a 12 años aproximadamente) y la etapa *formal*(a partir de los 12 o 13 años). Con las ingeniosas situaciones imaginadas por Piaget para demostrar su teoría, otros psicólogos han tratado de construir tests que permitieran evaluar las etapas del desarrollo del pensamiento infantil. Uno de los más conocidos es la Escala de Desarrollo del Pensamiento Lógico (EDLP) de Longeato(1969). En estos tests el problema no es como en los tests usuales comparar el rendimiento de un sujeto con el que según las normas tienen los niños de su misma edad, sexo y nivel socioeconómico, sino haciendo referencia a un criterio teórico que es la etapa de desarrollo al que su modo de razonamiento corresponde. La existencia de un sustento teórico tan preciso evita la concepción empírica de la estandarización. La validación de la escala por otro lado, descansa en un análisis jerárquico. Los ítemes que se corresponden con las distintas etapas de razonamiento deben resolverse en el orden que establece la teoría y esta adecuación va de 0 cuando el orden del éxito no es de lo que sería por azar, a 1 cuando el orden corresponde exactamente a la teoría. En esta escala los índices jerárquicos de las distintas escalas están alrededor de .90 lo que indica que los ítemes correspondientes a las distintas etapas se ordenan de acuerdo a la teoría piagetiana. Sin embargo, estos tests no han tenido gran éxito pues se vio pronto que un mismo

sujeto podía caracterizarse por etapas muy diferentes en los distintos subtests que tiene la batería, lo que hacía que no fuera posible caracterizar un sujeto por su etapa de desarrollo cognitivo y además se veía que si se comparan los resultados con los tests usuales de C.I. (como el WISC) la correlación era muy alta pero aportaba menor información- Los análisis de correspondencia realizados con estos tests no revelan tampoco factores distintos a los hallados con los tests de aptitudes comunes (Lautrey, Rieben & de Ribaupierre, 1986).

Existen ya algunos tests como el de Kaufman para conocimientos (Kaufman, A. S & Kaufman, N. L., 1985) que informan al examinado sobre los errores que comete. En este test las escalas de procesos mentales se derivan de la teoría sobre el proceso neurofisiológico en la especialización cerebral de Sperry y la investigación neurológica de Luria (1965). Las escalas enfocan el proceso usado para resolver un problema diferenciado del proceso secuencial lineal y analítico propio del hemisferio cerebral izquierdo (en los diestros) del procesamiento simultáneo, holístico y giestáltico que corresponde al derecho. Así, atendiendo a *cómo* se realiza la solución del problema se pueden realizar intervenciones clínicas terapéuticas para mejorar el aprendizaje, si parece necesario.

En este interés por estudiar el *como* se logra la resolución de un problema más que en el resultado final escueto, ha sido pionera en nuestro país la investigación del Dr Horacio Rimoldi (1991) y de sus colaboradores del CIIPME. Nosotros en un trabajo recientemente publicado (Cortada de Kohan, N. & Gentile, S. 1997) hemos tenido oportunidad de verificar lo importante que podría ser para los alumnos conocer los errores que cometen, al analizar de manera sistemática los errores que cometieron una muestra de alumnos secundarios en los ítems del Test de Raven (Forma Avanzada). Estudiamos los errores que cometían por haber fallado en una sola de las variables que era necesario analizar para llegar a solucionar cada ítem de la prueba y dividimos los errores en *buenos* y *malos*. Llamamos buenos a los errores en los que los sujetos habían fallado en una sola variable para llegar a la solución correcta y malos los errores en los que habían fallado en dos o más variables. Luego dividimos la muestra de los alumnos ($n = 165$) en el 27% que habían obtenido los puntajes totales más elevados y el 27% que habían obtenido los puntajes más bajos y establecimos con la prueba de χ

2 que existía una dependencia indudable ($p < .01$). Los alumnos mejores en sus resultados totales no solo cometían menor cantidad de errores en total, sino que aquellos que cometían eran casi siempre del tipo de errores buenos.

Algunos autores (Howell, 1986; Vygotsky, 1978) consideran que deben establecerse medidas directas del aprendizaje llamadas procedimientos de evaluación dinámicos que se refieren a una serie de procedimientos de evaluación dirigidos directamente a identificar y remediar problemas de aprendizaje. Esto puede significar un muestreo múltiple del aprendizaje de los niños, diferente de las evaluaciones usuales, que solo se realizan una o dos veces al año. Estas evaluaciones significan una estimación más confiable de la actuación de los jóvenes. Meyers, Pfeiffer y Erlbaum (1985) sugieren que la evaluación debe medir la interacción de la tarea del aprendizaje y el ambiente. En general esta evaluación de procesos surge de la psicología experimental en los cuales el análisis cognitivo de las tareas y el análisis de los errores, son importantes.

También conviene hacer referencia de un test interesante en este sentido como el de Woodcock-Johnson, cuya versión en castellano es Woodcock-Muñoz (1996) que se basa en el procesamiento intelectual de Horn (1966) llamada *Teoría Gf-Gc*. *Gf* sería *inteligencia fluida* y es la aptitud para resolver problemas nuevos que no

dependen de la educación y la cultura de los examinados y G_c es la *inteligencia cristalizada* que es el conocimiento y aptitudes de una persona, relacionados con su nivel educativo y el medio cultural en que vive. La teoría de Horn, inspirado por Cattell, pasó de las dos inteligencias a una teoría de las *inteligencias* pues el funcionamiento intelectual sería una estructura similar a la Vía Láctea. No se sabe cuantas inteligencias hay pero gracias al A.F. podemos establecer constelaciones diferenciando aptitudes amplias y aptitudes específicas. Hasta ahora se habrían determinado: Memoria a corto plazo (G_{sm}), Aptitud cuantitativa (G_Q), Procesamiento verbal (G_v), Procesamiento auditivo (G_a), Almacenamiento o recuperación asociativa (G_{ir}), Razonamiento fluido (G_f), Velocidad de procesamiento (G_s), y T tiempo de decisión (G_t).

Nuestra reflexión es que actualmente resulta importantísimo para educadores y psicólogos informarse sobre los procedimientos para evaluar y esto no solo desde el punto de vista psicométrico procurando tomar los exámenes de tipo lo más objetivo posible, lo cual es una condición *sine qua non*, sino también sobre las nuevas técnicas de evaluación experimental y cognitiva que proporcionan una devolución al alumno para mejorar el aprendizaje.

La psicología actual considera beneficioso evaluar los procesos con un optimismo realista (Schneider, 2001) y apreciar con una perspectiva más positiva las fortalezas y los logros para poder crear las oportunidades que mejoren la experiencia vital de las personas.

Citas bibliográficas

Alonso, M. M. (1994). Los psicólogos en la Argentina. Datos cuantitativos. *Acta Psiquiátrica y Psicológica de América Latina*. 40 (1) 50-55.

Anastasi, A. (1967). Psychology, psychologists and psychological testing. *American Psychologist* 22, pp297-306.

Carroll, J. B., (1976). Psychometric tests as cognitive tasks, E. Resnik (Eds) *The Nature of intelligence*. Hillsdale, Erlbaum

Carroll J. B., (1993). *Human cognitive abilities, a survey of factor analytic studies*. Cambridge, Mass. Harvard Univ. Press.

Cohen J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, Hillsdale N.J. Lawrence Erlbaum Pub.

Cortada de Kohan, N. (1998). La teoría de la Respuesta al Item y su aplicación al test verbal Buenos Aires. *Interdisciplinaria Vol 15, 1 y 2*.

Cortada de Kohan, N. (1999). *Teorías psicométricas y Construcción de Tests*. Buenos Aires, Edit. Lugar.

Cortada de Kohan, N. (2000). *Técnicas Psicológicas de Evaluación y Exploración*, México, Trillas.

Cortada de Kohan, N. y Gentile, S. (1997). Test de las Matrices Progresivas de Raven (Forma Avanzada) en jóvenes de 17 años. *Interdisciplinaria*, 14, 1-2, pp 107-122.

Cronbach, L. J., Gleser, G. C., Nanda, H. y Rajaratman (1972). *The dependability of behavioral measurements*. New York, John Wiley

Dawes, R. M. (1988). *Rational choice in an uncertain world*. San Diego, Ca. Hartcourt and Brace.

Foulquié, P. (1967). *Diccionario del Lenguaje Filosófico*. Barcelona, Labor.

Galibert, M. S., Aguerri, M. E y Attorresi (2000). Pesos óptimos de los ítemes en la elaboración de puntajes. *Revista Latinoamericana de Psicología*, Vol 32, N.1

Glaser R.(1963) Instructional technology. And the measurement of learning outcomes.. *American Psychologist* Vol18, pp. 519-543.

Glaser R(1981) Instructional Psychology, Past, Present and Future. *American Psychologist* Vol37, N3 pp292-305.

Goetz, E.T., Hall, R.J. y Fresco G (1990) Implications of cognitive Psychology for Assessment of Academic skills En C.Reynolds and R.W.Kamphaus (Edit) *Handbook of Psychological and Educational Assessment of Children* N. Y. The Guilford Press.

Grimm, L.G. y Yarnold, P. R. (1997) *Reading and Understanding Multivariate Statistics*, Washington, D. C. American Psychological Association.

Horn J. L. (1966) Refinement of the Theory of fluid and crystalized intelligence *Journal of Educational Psychology*, 57, 253-270.

Howell, K.W. (1986) Direct assessment of academic performance. *School Psychological Review*, 15, 324-335.

Hunt, E. B. (1971). What kind of computer is man? *Cognitive Psychology*. 2, pp.57-98.

Hunt, E. B. (1980). Intelligence as an information processing concept. *British Journal of Psychology*. 71, pp. 449-474.

Huteau, M y Lautrey, J.(1997). *Les Tests d'Intelligence*, Paris, La De-courante.

Jonson, D. E. (2000). *Métodos multivariados aplicados al análisis de datos*. México, International Thomson Editores.

Kaufman A.S. y Kaufman, N.L.(1985) *Kaufman Tests of Educational Achievement, Comprehensive Form Manual*, Circle Pine MN, American Guidance Service

Koch, S (1981) The Nature and Limits of Psychological Knowledge, Vol 36, N 3 pp257-269.

Lautrey, J. Rieben L. y Ribanpierre A. Edit. (1986). Les differences dans la forme du development cognitiv évalué avec des epreuves piagetienes: une application de l'analyse des correspondences. *Cahiers de psychologie cognitive* 6, 575-630.

- Longeot, F. (1969) *Psychologie differentielle et theorie operatoire de l'intelligence*. Paris, Dunot.
- Lord, F.M.(1980). *Applications of Item Response Theory to Practical Testing Problems*. N.J. Lawrence Erlbaum Ass.Publ.
- Luria,A.R.(1961) Study of the abnormal child. *American Journal of Orthopsychiatry*, 31,1-10.
- Meehl, P.E.(1954) *Clinical versus statistical prediction*, Minneapolis, University of Minnesota Press.
- Meyers, J., Pfeiffer, J. y Erlbaum, P. (1985) Process assessment. A model for broadening assessment. *Journal of Special Education*, 19, 73 y sigs.
- Miller, G.(1967) *The Psychology of Communication*. Londres, Penguin.
- Neisser, U(1967) *Cognitive psychology* N.Y. Appleton Century Crofts.
- Newell, A (1973) You can't play twenty questions with the nature and win. En Chase, R (Edit) *Visual information Processing*, N.Y. Academic Press.
- Newell A. and Simon H. A.(1972). *Human Problem Solving*. N. J., Prentice Hall.
- Piaget, J (1970). *L'Epistemologie genetique*. Paris, PUF.
- Pinker, S (1994). *The Language Instinct*. N. Y., Harper Collins.
- Pinker, S (2001). *Cómo funciona la mente* Bs. As. Planeta.
- Rasch,G.(1980). *Probabilistic Models for some Intelligence and Attainment Tests* (Reprint) Chicago, University of Chicago Press.
- Rimoldi, H.(1991). Solución de problemas. Problemas que trae un problema. En: Mayor, J. y Pinillos, J. L. (Eds). *Tratado de Psicología General Vol 5*. Madrid, Alhambra Longmans.
- Sampson, E. E. (1981). Cognitive Psychology as Ideology. *American Psychologist* Vol36, N 7 pp 730-743.
- Schneider, S. L. (2000) In search of realistic optimism. *American Psychologist* Vol 56 - N3.
- Sternberg, R. J.(1981). Testing and cognitive psychology *American Psychologist* Vol 36 N.7, pp. 470.
- Sternberg, R. J.(1987). (Eds) *La Inteligencia humana*, Barcelona, Paidós.
- Vygotsky, A.S. (1978). *Mind in Society. The Development of higher Psychological Processes*. Cambridge, Mass. Harvard Univ. Press.
- Weinstein, C., Goetz, E. and Alexander, P.A.(1988). *Learning and Study strategies* San Diego CA. Hartcut and Brace, Co.

Wittrock, M. C. (1990). Generative Processes of comprehension *Educational Psychology*, 24 pp 345-376.

Wittrock, M. C. y Baker, E. L. (1989). *Testing and cognition* N. J. Prentice Hall.

Woodcok-Muñoz (1996). *Batería R. Pruebas de habilidad cognitiva y pruebas de aprovechamiento*, Itaka, Riverside Publishers.



USAL
UNIVERSIDAD
DEL SALVADOR