

**Se autoriza el uso de este material citando su procedencia:  
Noblejas de la Flor, M.A. (2000). Fiabilidad de los tests PIL y Logotest. NOUS: Boletín de Logoterapia y Análisis Existencial. (4), 81-90.**

*FIABILIDAD DE LOS TESTS PIL Y LOGOTEST*

**M<sup>a</sup> Ángeles NOBLEJAS DE LA FLOR**

## **INTRODUCCIÓN**

Validar instrumentos o crear nuevas herramientas que ayuden en la praxis logoterapéutica es necesario para ubicar nuestro quehacer en el actual paradigma científico, en pos de su objetividad.

La ciencia es, al mismo tiempo, capaz de progresar y falible. Por ello, contando con una postura abierta, ha de reconocer sus límites, desde una actitud ética. En esta línea, suscribimos las palabras de Boado de Landaboure (1994, p. 1):

“La hiper o hipovaloración de test y técnicas, encubre muchas veces la inseguridad personal de quien maneja el recurso y conlleva a la empobrecida circunstancia de trastocar lo que es propiamente humano por el resultado de lo técnico. Esto, sin duda, ha engendrado distorsiones iatrogénicas originando fundadas críticas desde diversos enfoques. Estas críticas no invalidan los instrumentos si no que plantean la necesidad de conocer sus alcances y límites.

Tanto la rutina profesional como la innovación apresurada entrañan el grave riesgo de la despersonalización”.

Para avanzar en el conocimiento del hombre estudiamos aspectos particulares, pero nunca hemos de olvidar que el ser humano es siempre más de lo que sabe de sí y puede saber más de lo que algún otro sabe de él (Jaspers, 1996).

Finalmente, debemos resaltar la necesidad de asumir lo que otras disciplinas, como pueden ser en este caso la estadística, han establecido como procedimientos objetivos; siempre que no nos sitúen en una visión reduccionista y determinista del ser humano.

En esta línea, en este artículo nos planteamos estudiar la fiabilidad de la versión española de los dos test logoterapéuticos más difundidos: el Purpose In Life (PIL) desarrollado en los Estados Unidos (Crumbaugh, 1969) y el Logotest desarrollado en Alemania (Lukas, 1986). Los lectores interesados pueden encontrar una breve descripción de ambos test en un trabajo anterior (Noblejas, 1999).

Dos son los requisitos fundamentales que han de cumplir los instrumentos de medición psicológicos para asegurar la calidad de los datos que serán analizados: la validez y la fiabilidad.

La validez se puede definir (Bisquerra, 1987) como el grado en que un instrumento mide lo que pretende medir. Esta afirmación parece algo obvio pero, por ello, es quizás el aspecto más complicado de objetivar. Se habla de varios tipos de validez ligados a procedimientos específicos de tratamiento del problema. La ‘validez de contenido’ se da cuando los items representan el contenido que se pretende medir. En este sentido, desde los conocimientos de la logoteoría se puede corroborar que los items representan los contenidos relacionados con la cuestión logoterapéutica del sentido; específicamente Lukas explicita el objetivo de cada una de sus preguntas y su conexión con la logoteoría. Este último aspecto nos llevaría directamente a otro concepto muy cercano que sería la ‘validez de constructo’, que determina en qué medida

el test es congruente con una teoría o un constructo hipotético. Hay varias formas de objetivar esta validez, una de las más habituales es el análisis factorial con el que se confirma la estructura que se asigna a un test. En este sentido, podemos recordar el ajuste de las partes del logotest a sus componentes factoriales (Noblejas, 1999).

La 'validez concurrente' consiste en comparar las puntuaciones de la prueba con otros indicadores paralelos o criterios externos. En este sentido se han realizado diferentes trabajos como comparar las puntuaciones de sentido con pruebas de psicohigiene general (Lukas, 1986), ansiedad, autoestima, satisfacción en la vida, bienestar existencial (Shek, 1993), e incluso con el test cumplimentado por personas cercanas a las que lo rellenaban inicialmente (Crumbaugh, 1969).

Por último, la 'validez predictiva' ofrece un índice del grado en que una prueba sirve para predecir comportamientos futuros. En este sentido Crumbaugh (1969) fue capaz de predecir el orden de puntuaciones de varios grupos con ciertas características y Noblejas (1997 y 1998) el comportamiento de grupos de personas con problemas de drogadicción en programas de tratamiento.

La fiabilidad se puede definir (Bisquerra, 1987) como la constancia o estabilidad de los resultados que proporciona un instrumento de medida. Los coeficientes de fiabilidad se interpretan como una correlación, considerándose que valores superiores a 0,75 indican alta fiabilidad. La objetivación de los diferentes matices de los aspectos recogidos en el test en diversas preguntas conducen a una mayor fiabilidad y al mismo tiempo al aumento del número de items planteados. Esto aclara la regla que indica que a más items, mayor fiabilidad y que, incluso, nos permite calcular cuántos items necesitaríamos para alcanzar una fiabilidad deseada.

Hay varios métodos para evaluar la fiabilidad:

- Métodos de consistencia interna. Se basan en obtener coeficientes de consistencia de los datos recogidos en los diferentes tests rellenados. Los más empleados son los coeficientes de Kuder-Richardson (específicamente el que se conoce como fórmula 20) que se usan para tests con items dicotómicos y el coeficiente alfa de Cronbach, aplicable a escalas de varios valores posibles (y que coincide con el de Kuder-Richardson cuando se usa en su cálculo la varianza descriptiva en lugar de la estimada).
- Método del test-retest. Se basa en aplicar la prueba en dos ocasiones distintas a las mismos sujetos y evaluar la correlación entre ambas mediciones. Tiene el inconveniente de que nada asegura cambios en la situación de las personas entre una prueba y otra, variable que habrá que intentar controlar en lo posible.
- Método de formas paralelas. Se administran dos formularios similares, que no son iguales pero que intentan medir lo mismo, a un grupo de personas y se estudian la similitud de los resultados por medio de la correlación, pero con algunas correcciones de sesgo.
- Método de las dos mitades. Consiste en dividir el test en dos partes (pares frente a impares o la primera mitad frente a la segunda) y buscar la correlación entre ambas (para lo cual se puede usar el coeficiente de Spearman-Brown o los coeficientes de Guttman).

Estos métodos están tomados de la Teoría Clásica de los Test; no abordamos la aproximación de la Teoría de la Generabilidad o la de la Respuesta al Item (se puede consultar Martínez Arias, 1996) por estar considerando un test de autoinforme sencillo sin que se hayan considerado factores de variación como observadores, situaciones, ensayos, grupos de tareas, etc. Sin embargo, en futuros trabajos, se podría tener en consideración la Teoría de la Generabilidad para la puntuación de la parte libre del Logotest por varios observadores independientes.

## **MATERIAL Y MÉTODO**

Los tests PIL y Logotest fueron rellenados por 841 personas en un muestreo no probabilístico de tipo casual.

La descripción de ambos instrumentos puede encontrarse en sus manuales (Crumbaugh y Maholick, 1969; Lukas, 1986) o en una síntesis presentada anteriormente (Noblejas, 1999). Para facilitar la lectura del presente artículo incluimos la agrupación de preguntas en factores y partes en las tablas 1 y 2.

**Tabla 1. Agrupación factorial de las preguntas del test PIL.**

FACTORES	PREGUNTAS (por orden de importancia en el factor)
Experiencia de sentido (FP2)	P2 P5 P1 P9 P19 P20 P17
Percepción de sentido (FP1)	P11 P16 P4 P6 P10 P12 P9 P20 P17
Metas y tareas (FP3)	P13 P3 P8 P7 P20 P17 P19
Dialéctica destino-libertad (FP4)	P14 P15 P18

**Tabla 2. Agrupación en partes y factorial de las preguntas del Logotest.**

PREGUNTAS	PARTES del test	FACTORES					
		FL1	FL2	FL3	FL4	FL5	FL6
LI1	Parte I Contenidos de sentido						X
LI2				X			
LI3							X
LI4				X			
LI5							X
LI6					X		
LI7					X		
LI8					X		
LI9			X				
LII1	Parte II Síntomas de frustración existencial	X					
LII2		X					
LII3		X					
LII4		X					
LII5				X	X		
LII6		X					
LII7		X					
LIII1	Parte III					X	
LIII2	Casos propuestos y propio					X	
LIII3A Escala de Sentido			X				
LIII3B Escala de Actitud			X				

Los nombres de los factores son: FL1 Frustración existencial, FL2 Balance existencial, FL3 Metas y tareas, FL4 Intereses y compromisos, FL5 Flexibilidad axiológica y FL6 Sentido en el entorno.

Como análisis estadísticos de la fiabilidad hemos seleccionado, de entre los más frecuentes en la bibliografía: el coeficiente alfa de Cronbach y los coeficientes de Guttman (en concreto recogemos el Lambda 4 o de 'límites inferiores' y el Lambda 6) y de Spearman-Brown para el análisis en dos mitades (por el método de preguntas pares frente a impares).

Todos los análisis se han realizado con el procedimiento RELIABILITY del paquete de estadística SPSS.

## RESULTADOS Y DISCUSIÓN

Los resultados encontrados para los dos test como conjuntos los recogemos en la tabla 3.

**Tabla 3. Estadísticos de fiabilidad de los test logoterapéuticos.**

<b>ESTADÍSTICO</b>	<b>Test PIL</b>	<b>Logotest</b>
Alfa	0.879	0.513
Alfa estandarizado	0.892	0.539
Guttman (Lambda 4)	0.838	0.412
Guttman (Lambda 6)	0.893	0.584
Spearman-Brown	0.838	0.591

Los estadísticos encontrados para el test PIL indican una gran fiabilidad, similar a la encontrada tanto por sus autores como por otros estudios: Shek (1986) en China (0.84 para la alfa de Cronbach y 0.82 para el coeficiente de Guttman) o Reker y Cousins (1979) en Estados Unidos (0.87 para Spearman-Brown y 0.79 para un test-retest).

Centrándonos, ahora, en el Logotest, podemos ver que los índices de fiabilidad son bajos. La autora del mismo (Lukas, 1986), informa en su baremación de un coeficiente de fiabilidad de Spearman-Brown de 0.835, para una división de la parte I del test en dos mitades, lo que supone un índice muy alto para un sub-test de 9 preguntas. Así mismo, justifica con la teoría de Rash la unidimensionalidad del test y la incorporación de items.

Las relaciones de las distintas partes del test con la estructura factorial y sus relaciones ya las estudiamos (Noblejas, 1999) con resultados muy similares respecto a los planteamientos de la autora en su construcción. Sin embargo, en el caso de la fiabilidad, la versión española no da resultados parecidos.

Intentamos, en todo caso, una aproximación más detallada a las diferentes partes del test, aunque sólo con el coeficiente alfa de Cronbach ya que permite el estudio de la fiabilidad con bajo número de preguntas (no hay que dividir las en dos grupos) y porque las variaciones entre los distintos índices de fiabilidad no son importantes. Para ello se han estudiado: la fiabilidad de las partes originales de la prueba, de combinaciones de éstas, así como de los factores y algunas combinaciones de éstos. Se obtienen los resultados de la tabla 4.

**Tabla 4. Estadísticos de fiabilidad para algunas partes del Logotest.**

<b>LOGOTEST agrupaciones</b>	<b>Coef. Alfa</b>	<b>Coef. Alfa estandarizado</b>
Total	0.513	0.539
Parte I	0.364	0.367
Parte II	0.507	0.527

Parte III	0.335	0.353
Partes I y II	0.459	0.493
Partes II y III	0.486	0.512
Factor 1 ( $\approx$ Parte II)	0.289	0.418
Factor 2 (LIII3 A y B)	0.635	0.687
Factor 3	0.342	0.332
Factor 4	0.400	0.400
Factor 5 (LIII 1 y 2)	0.604	0.604
Factor 6	0.299	0.307
Factor 3, 4 y 6 ( $\approx$ Parte I)	0.418	0.408

Para la definición de factores ver Noblejas (1999), aunque aquí se han considerado las preguntas sin coeficientes de ponderación.

Según estos datos, para la versión española del Logotest, la parte de la estructura original que más fiabilidad aporta sería la II o recogida de los síntomas de frustración, que cuenta con 7 preguntas. La parte I (que recoge las áreas de sentido con 9 preguntas) y la parte III (preguntas por la felicidad –LIII1-, sufrimiento –LIII2- y escalas de sentido –LIII3A- y actitud –LIII3B-), son las que aparecen con peor fiabilidad. Esta última parte III, sin embargo, analizada como los dos factores diferenciados que contiene (F2 y F5), aporta un importante incremento de la fiabilidad, máxime considerando que sólo están compuestas de dos puntajes cada uno, llegando a 0.60.

La consideración del resto de los factores no tiene mayor interés; simplemente resaltar que la corrección de la parte I y parte II en las dos preguntas que cambian de factor (ver Noblejas, 1999) hacen, respectivamente, aumentar y reducir su fiabilidad ligeramente.

Tenemos que destacar que hasta ahora no hemos tenido en cuenta uno de los aspectos que influye sobremedida en los resultados de fiabilidad: la longitud del test, o número de ítems que lo componen. Basándonos en la ecuación de Spearman-Brown, podemos obtener la fiabilidad de un test con un número distinto de ítems paralelos al test original (ver Martínez Arias, 1996):

$$C_f = N_f \cdot C_i / (N_i + (N_f - N_i) \cdot C_i)$$

Siendo  $C_f$  el coeficiente de fiabilidad para una longitud final  $N_f$  y  $C_i$  el coeficiente de fiabilidad de la longitud inicial  $N_i$ .

Aplicando esta corrección para un hipotético test (semejante al Logotest) de 20 preguntas, la parte II pasaría a tener una fiabilidad de 0.761 y la parte III, 1 y 2, (correspondiente al factor 5), tendría una fiabilidad de 0.813. Esto, considerando que se esté en disposición de desarrollar un test de 20 preguntas sobre los aspectos indicados sin problemas en las propiedades de los ítems, ya nos situaría en condiciones de fiabilidad más aceptables.

En definitiva, esta baja fiabilidad del Logotest puede estar ilustrando un viejo problema de la teoría clásica de los test. La fiabilidad (interna) exige altas correlaciones entre los ítems y la validez (predictiva) exige bajas correlaciones entre los mismos para atender a los diferentes aspectos del tema que se pretenden recoger. Recuérdese además que el Logotest pretendía abordar aspectos no presentes en el PIL y que tiene una estructura intrínseca más diferenciada.

En esta disyuntiva, muchos autores aconsejan (ver, por ejemplo, Guilford y Fruchter, 1978) que cuando se dan conflictos entre el interés de recoger aspectos diferenciados de un constructo y el deseo de obtener coeficientes de fiabilidad altos se opte por baterías de test y no por tests simples.

## CONCLUSIONES

El estudio de la fiabilidad de los tests PIL y Logotest, a partir de los datos obtenidos en las baremaciones españolas, nos lleva a verificar el comportamiento del PIL como una prueba fiable. Sin

embargo, se han encontrado unos coeficientes bajos para el Logotest. Esto último nos plantea la posibilidad de una mejora de este test, posiblemente a través de estudios y desarrollos estadísticos relacionados con el incremento del número de preguntas para cada uno de los factores que comprende.

*M<sup>a</sup> Ángeles NOBLEJAS DE LA FLOR es doctora en CC. de la Educación y Vicepresidenta de la Asociación Española de Logoterapia.*

## BIBLIOGRAFÍA

- Bisquerra, R. (1987). *Introducción a la estadística aplicada a la investigación educativa*. Barcelona: Promociones Publicaciones Universitarias.
- Boado de Landaboure, N.B. (1994). *Escala existencial de A. Längle y C. Orgler. Un instrumento técnico para el abordaje preventivo, diagnóstico y terapéutico*. Buenos Aires: Fundación Argentina de Logoterapia.
- Crumbaugh, J.C. y Maholick, L.T. (1969). *Manual of instructions for the Purpose In Life test*. Saratoga: Viktor Frankl Institute for Logotherapy.
- Guilford, J.P. y Fruchter, B. (1978). *Estadística aplicada a la psicología y la educación*. Bogotá: McGraw Hill.
- Jaspers, K. (1996). *Psicopatología general*. México: Fondo de Cultura Económica.
- Lukas, E. 1986 (Versión castellana de 1996) *Logo-test. Test para la medición de la 'realización interior de sentido' y de la 'frustración existencial'*. Ed. Almagesto: Buenos Aires. (Se ha confrontado también con otra traducción del original alemán de J.Coloma).
- Martínez Arias, R. (1996). *Psicometría: teoría de los tests psicológicos y educativos*. Madrid: Síntesis.
- Noblejas, M.A. (1997). Meaning levels and drug-abuse therapy: an empirical study. *The International Foun for Logoterapy*, 20, (1), 46-52.
- Noblejas, M.A. (1998). Evaluación de los abandonos en las primeras fases de un programa de tratamiento de drogadicción. *NOUS*, (2), 73-79.
- Noblejas, M.A. (1999). Estructura factorial de los tests PIL y Logotest. *NOUS*, (3), 67-84.
- Reker, G.T. y Cousins, J.B. 1979. Reliability of the Seeking Of Noetic Goals (SONG) and Purpose In Life (PIL) tests. *Journal of Clinical Psychology*. 35, (1), 85-91.
- Shek, D.T.L. 1986. The purpose in life questionnaire in a chinese context: some psychometric and normative data. *Chinese Journal of Psychology*. 28, (1), 51-60.
- Shek, D.T.L. (1993). The chinese Purpose-In-Life test and psychological well-being in chinese college students. *The International Forum for Logotherapy*, 16, (1), 35-41.