

LOS TESTS DE ESP, ¿POSEEN CONFIABILIDAD?*

J. Ricardo Musso

INTRODUCCION

Este ensayo nos fue sugerido por un trabajo de Eysenck (1967) en el que recomienda a los parapsicólogos la realización de estudios sobre la confiabilidad de los puntajes que obtienen en sus tests. Los juzga particularmente importantes para los estudios sobre la dependencia de la ESP respecto de determinadas variables; por ejemplo, los rasgos de personalidad, los efectos de posición, etc. Dice así:

"En la construcción de un test de personalidad o habilidad, el psicólogo pone gran énfasis en la confiabilidad de sus instrumentos, es decir, investiga el grado en que sus sujetos producen resultados repetibles solo si la confiabilidad razonablemente alta vale la pena continuar estudiando la validez del instrumental, es decir, descubrir si mide lo que se supone que mide, o usarlo en estudios correlacionales. Si un test no se correlaciona consigo mismo, o sea, no suministra resultados reproducibles, sería absurdo esperar que se correlacione con otras mediciones; por ello una falla en esta última correlación no debería interpretarse como indicadora de la ausencia de una relación verdadera entre la habilidad que el test intenta (pero aparentemente no logra) medir o la otra medición (1967, p.64).¹

LAS RELACIONES ENTRE LA CONFIABILIDAD Y LA VALIDEZ

Consideramos que la recomendación de Eysenck es muy valiosa, pero nos interesa reflexionar sobre la idea que, al parecer, expresa en las partes que hemos subrayado en la cita anterior. La entendemos en el sentido de que si un test no suministra resultados repetibles, esto revela que el test no logra medir la habilidad que intenta medir (en otras palabras, que no tiene validez para el uso que se le destina). Es una idea que parece fundarse en la más ortodoxa teoría psicométrica según la cual "un test puede ser *confiable* aún sin ser una medición válida de lo que se intenta medir, pero un medición nunca es válida a menos que sea *confiable*." (McCollough & Van Atta, 1963, p.290). Según esta forma de ver las cosas, la confiabilidad, por lo tanto es una condición necesaria (aunque no suficiente) de la validez.

Los constructores de test psicológicos suelen realizar profundos estudios sobre la confiabilidad de sus tests, antes de investigar su validez para determinados usos. Pero en parapsicología estos estudios están notablemente ausentes. Numerosos tests (con cartas de ESP, con dibujos, etc.) se aplican en investigaciones de ESP; sin embargo, se desconoce su grado de confiabilidad.

Pensamos que esta aparente falla se debe a que el objeto de estudio de los parapsicólogos difiere de los objetos de estudio de los psicólogos (las aptitudes, rasgos, etc.) en una importante propiedad: la constancia de la propiedad inobservable medida, la invarianza de sus efectos observados, por los cuales se la mide.

LA CONSTANCIA DE LA PROPIEDAD Y LA CONFIABILIDAD DE SUS MEDICIONES

La constancia de la propiedad que se mide en un objeto es una característica presupuesta, en los estudios de confiabilidad. En general, si se comprueba que los puntajes de un test no son estables no se correlacionan entre si se dice que el test carece de confiabilidad. Lo que expresa esta declaración no es que la propiedad medida repetidas veces en los sujetos muestra variaciones de medición en medición, sino que esta se ha mantenido constante, pero el test falla en medirla. "La confiabilidad -dice Guilford (1956, p.436)- es de un cierto instrumento aplicado a cierta población bajo ciertas condiciones. Y una de esas condiciones es que la propiedad medida sea constante. Lo son en un grado apreciables las propiedades que miden los tests psicométricos. Si no lo fueran, los puntajes de diferentes mediciones de una misma propiedad efectuadas con ellos a un mismo grupo de individuos, no presentarían el alto grado de correlación consigo mismo que generalmente muestran, y por el cual se afirma que son confiables.

Pero la ESP no es constante. Rhine y otros han señalado que la *erraticidad* es una de las características de la ESP comprobada experimentalmente (Rhine *et al.*, 1966, p.313), y en un reciente experimento (Musso & Granero, 1979, 1981), hemos podido corroborar la verdad de este aserto. Sujetos que en ocasiones obtienen en tests de ESP puntajes que obligan a admitir que su ESP estuvo muy activa, en otras ocasiones sólo suministran puntajes de azar, sin que hasta ahorase haya

podido formular criterios para predecir cuando estara activa y cuando no. Además, en ocasiones la ESP funciona en forma positiva (psi hitting) y en otras ocasiones en forma negativa (psi missing), y tampoco hay criterios para predecir en que dirección (+ ó -) operará.

Por ello, parece claro que mientras no se logre identificar un conjunto de condiciones en el cual la ESP de los sujetos funcione, en forma previsible y asegurar que ellas esten presentes en la oportunidad en que se aplican los tests para medir su ESP, no es de esperar que esos tests proporcionen puntajes, resulten repetible.

LA VALIDEZ PRESUPONE A LA CONFIABILIDAD

¿Significa la falta de correlación consigo mismo de los tests de ESP que éstos carecen de confiabilidad? Si tal cosa se aceptara entonces cabría admitir que no tienen validez. Este sería un corolario de la premisa antes referida: la de que "una medición nunca es válida a menos que sea confiable". Pero nosotros preferimos apoyarnos en esa premisa para afirmar que los tests de ESP poseen confiabilidad.

Las investigaciones sobre la existencia de la ESP satisfacían el requisito que Guilford (1954, p.898) considera fundamental en los estudios sobre la validez de un test: la prueba de la correlación de un puntaje con un criterio externo (el criterio externo es el objetivo, con el cual se correlaciona la respuesta) fueron claramente satisfechas, se obtuvieron resultados extra-azar, al aplicarse los tests de significación estadística. Estos últimos tests prueban que un factor causal estuvo activo en la determinación de esa correlación y el control de las condiciones experimentales prueba que ese factor *no pudo ser otro* que la ESP. Con ello se prueba que esos tests tienen validez para medir la ESP. No es una validez teórica, sino comprobada en las aplicaciones prácticas, experimentales. En consecuencia, esos tests deben poseer confiabilidad, pues esta propiedad -según vimos antes- es una condición *necesaria* de la validez según la teoría psicométrica. Deben poseer confiabilidad, aunque todavía esta ni puede medirse por la técnica ejemplificada por Eysenck, que se denomina correlación test-retests.

LOS PUNTAJES EXTRA-AZAR Y LA CONFIABILIDAD DE LOS TESTS

Eysenck parece ser contrario a la idea que acabamos de exponer. "A veces -dice- se ha arguido que la demostración de resultados extra-azar prueba la confiabilidad del test, pero no es así" (Eysenck, 1967, p.65). Y en apoyo de su opinión presenta una tabla que muestra los puntajes (imaginarios) de 12 sujetos que hicieron cada uno de los dos juegos de GESP con cartas de ESP. La tabla es la siguiente:

TABLA 1

SUJETOS	A	B	C	D	E	F	G	H	I	J	K	L	Totales	
Juego	1	5	4	6	9	8	5	7	9	15	11	14	10	103
Juego	2	14	11	15	7	9	8	5	6	9	4	5	14	107

Eysenck señala, correctamente, que la Tabla 1 muestra que los sujetos "obtienen puntajes bien por encima del azar (el total esperado por azar es $5 \times 12 = 60$ en cada juego), pero fallan completamente en mostrar cualquier evidencia de confiabilidad: una persona con alto puntaje en el juego 1 no necesariamente obtiene un puntaje alto en el juego 2, y una persona con puntaje bajo en el juego 1 no necesariamente obtiene un puntaje bajo en el juego 2 (Eysenck, 1967, p.65).

Calculamos la correlación interclásica de los dos juegos mediante el coeficiente de Pearson y determinamos un valor $r = -0,42$, 10 gl.($p = n.s.$). Claramente los puntajes, como dice Eysenck, no son lo consistentemente altos o bajos entre los juegos (Mas bien tienden a ser de signo inverso-correlación negativa). Pero, ¿significa este hecho que esos puntajes "fallan completamente en mostrar cualquier evidencia de confiabilidad"?

Si aceptáramos esto tendríamos que admitir, por lo que dijimos antes, que no tienen validez para medir la ESP, lo cual es claramente un absurdo porque la validez para medir la ESP en condiciones experimentales la hemos admitido por principio, y el resultado extra-azar es claramente significativo. El método que propone Eysenck supone la estabilidad de la ESP en el tiempo, y qua éste es la misma para todos los individuos.

*Grupo Rosarino de Estudios Psicológicos
Gálvez 2150
(2000) Rosario
SANTA FE*

REFERENCIAS

- CHILD, I.L. (1977). Statistical regression artifact in parapsychology. *Journal of Parapsychology* 41, pp.23-33.-
- CHILD, I.L. (1978). Statistical regression artifact: Can it be made clear? *Journal of Parapsychology* 42, pp.179-193.-
- EYSENCK, H.J. (1967). Personality and ESP. *Journal of the Society for Psychical Research* 61, pp.55-70.-
- GUILFORD, J.P. (1954). *Psychometric Methods*. McGraw Hill: New York, NY.
- GUILFORD, J.P. (1956). *Fundamental Statistics in Psychology and Education*. McGraw Hill: New York, NY.
- HAGGOD, M.J. & PRICE, D.O. (1955). *Estadística para Sociólogos*. Instituto de Intercambio de Estadística: Rosario (Santa Fe).
- MCCOLLOUGH, C. & VAN ATTA, L. (1963). *Statistical Concepts*. McGraw Hill: New York, NY.
- MCNEMAR, Q. (1955). *Psychological Statistics*. Wiley, New York, NY.
- MUSSO, J.R. GRANERO, M. (1979). *Análisis metodológico de una crítica a un experimento nuestro de ESP sobre dibujos ocultos*. Comunicación No.9. Instituto Argentino de Parapsicología: Buenos Aires.
- MUSSO, J.R. GRANERO, M. (1981). U-effects in an ESP experiment with concealed drawings. *Journal of Parapsychology* 45, pp.99-120.-
- RHINE, J.B.; PRATT, J.G.; SMITH, B.M.; STUART, C.E.; & GREENWOOD, J.A. (1966). *Extrasensory Perception After Sixty Years*. B.Humphries Ed.: Boston
- WINER, B.J. (1962). *Statistical Principles in Experimental Design*. McGraw Hill: New York, NY.

1- El subrayado es nuestro.

2- Ulterior información puede ser estudiada en los trabajos de Child (1977, 1978), Guilford (1956), Haggod & Price (1955), McNemar (1955), y Winer (1962).