

## LA CALIBRACIÓN DE JUICIOS PSICOFÍSICOS: ESTIMACIÓN DE MAGNITUDES

Ana Julia Garriga Trillo, Angel Villarino, María José González Labra y  
María Angeles Arnau

Facultad de Psicología. UNED, Madrid

La calibración ha sido ampliamente estudiada en tareas de juicios bajo incertidumbre. En la mayoría de estos estudios los sujetos muestran sobreconfianza en sus juicios. La calibración de juicios psicofísicos ha sido calculada en tareas de discriminación obteniéndose un efecto de subconfianza (Björkman, Juslin y Winnan, 1993). Los juicios de estimación de magnitudes no han podido ser calibrados pues una cuantificación de la respuesta ante un estímulo no se puede considerar directamente como acertada o no. Presentamos en este trabajo una forma de obtener la proporción de aciertos utilizando una medida ordinal aplicada a dos variantes de la técnica de estimación de magnitudes. Calculando la calibración, encontramos que en las tareas de estimación de magnitudes los sujetos manifiestan subconfianza, no existiendo diferencias entre técnicas ni relación con el grado de ajuste individual a la función de Stevens, su pendiente, su intercepto y el tiempo medio de reacción.

*Calibration of psychophysical judgments: Magnitude estimates.* Calibration has been widely studied in judgments under uncertainty. Most of these studies show a clear overconfidence effect. In psychophysical judgments calibration has been studied in discrimination tasks and a clear underconfidence effect has been found (Björkman, Juslin y Winnan, 1993). Magnitude estimation judgments have not been calibrated since a quantitative response cannot be directly considered as right or wrong. This work presents a way to measure the proportion of correct responses, at an ordinal level of measurement, using two variants of the magnitude estimation technique. Calculating calibration we found that in both magnitude estimation tasks subjects show an underconfidence effect and no relationship is found between this measure and the goodness of fit measure obtained from Stevens' function, its slope, intercept and mean reaction time.

En la investigación sobre los juicios bajo situaciones de incertidumbre en las cuales existe una base de conocimientos incompleta cabe destacar que uno de los aspectos relevantes de este proceso es la confianza que muestran los sujetos sobre sus propias evaluaciones. Es habitual estudiar los juicios de confianza por medio del análisis de la calibración (CA). Esta se define como la dife-

rencia entre la proporción de respuestas correctas (PC) y el grado de confianza media del sujeto sobre los juicios emitidos (CO), es decir:

$$CA = CO - PC \quad (1)$$

De esta forma la Ecuación 1 permitiría tres opciones posibles:

Opción I  $CA > 0 \rightarrow CO > PC$  (2)

Opción II  $CA < 0 \rightarrow CO < PC$  (3)

Opción III  $CA = 0 \rightarrow CO = PC$  (4)

En (2) se dice que el sujeto manifiesta una sobreconfianza en sus juicios, en (3) una

subconfianza y en (4) que el sujeto está calibrado. La calibración óptima dependerá de las expectativas realistas que muestren los sujetos al enfrentarse con distintas claves ambientales. Esta medida se utiliza muy a menudo como variable dependiente en la comparación de grupos y de sujetos (Glaser, 1989; Yates, 1990).

En estudios cognitivos clásicos los trabajos sobre calibración han puesto de manifiesto, en general, que los sujetos muestran un exceso de confianza al evaluar sus respuestas ante distintos tipos de situaciones (Cooper, Woo & Dunkelberg, 1988; Lichtenstein & Fischhoff, 1977; Lusted, 1977; Oskamp, 1965; Wagenaar & Keren, 1986). La robustez de este efecto viene descrita claramente por Lichtenstein, Fischhoff y Phillips (1982) al revisar, de forma casi exhaustiva, las investigaciones en las que se pretendía eliminar la sobreconfianza. Los autores concluyen esta revisión admitiendo la persistencia de este efecto a pesar de los distintos y variados esfuerzos metodológicos encaminados a su corrección. Este efecto de sobreconfianza parece ser otro sesgo sistemático más de la cognición humana.

Otro de los aspectos relevantes sobre juicios bajo incertidumbre, señalada por Howell (1971), es la distinción entre juicios bajo incertidumbre interna o adquirida y juicios bajo incertidumbre externa o dependiente del medio. En este último caso la incertidumbre se encontraría en el mundo externo, mientras que en el primer caso, posiblemente el de la mayoría de los juicios cognitivos tales como los emitidos en tareas de conocimiento general o los juicios expertos, la incertidumbre sería inherente al evaluador. Los resultados obtenidos por Howell podrían sugerir que la sobreconfianza surge de un tipo de incertidumbre y la subconfianza del otro.

En los juicios psicofísicos se dispone de una base de conocimientos incompleta y, por ello, claramente son juicios bajo incertidumbre (Garriga Trillo y Dorn, 1991) aunque los estudios sobre la evaluación de la confianza

y la calibración han sido escasos. Estos aspectos sólo han sido estudiados en tareas de discriminación sensorial, ya desde 1884, según citan Björkman, Juslin y Winman (1993). Estos autores encuentran que en la discriminación sensorial y utilizando el método de los estímulos constantes, el sujeto manifiesta subconfianza. Afirman también que este sesgo es, en la práctica, imposible de evitar y no se ve influido ni por la experiencia, ni por el nivel de dificultad de la tarea. Sin embargo, Skotnikova (1993) encuentra que utilizando una técnica de discriminación no ordenada, los sujetos manifiestan el sesgo de sobreconfianza. Considerando todos los estudios mencionados resulta evidente que el manifestar sobre o subconfianza es función de la tarea y/o de la técnica utilizada. No obstante, este último aspecto, dentro de la psicofísica, sólo se ha estudiado en tareas de discriminación sensorial.

En los juicios psicofísicos de estimación de magnitudes, además de una discriminación, al sujeto se le pide que emita un juicio numérico. Este tipo de tarea, dada su definición, conlleva procesos que podemos considerar «más» cognitivos. La calibración en este tipo de juicio psicofísico no ha sido estudiada. Si pudiésemos calcular esta medida podríamos comparar el efecto de sobre y subconfianza en función del tipo de juicio (sensorial o cognitivo) y/o de la técnica utilizada, como en los estudios de Glaser (1989) y Yates (1990) para comparar distintas muestras. Por otra parte, cabría la posibilidad de estudiar la relación de esta medida, y sus componentes, con variables resultantes del escalamiento psicofísico derivadas de las técnicas de estimación de magnitudes. Algunas de las variables resultantes del escalamiento podrían ser el coeficiente de determinación ( $r^2$ ), la pendiente y la ordenada de la recta obtenida en el ajuste escogido para representar la relación entre las magnitudes estimulares, E, y las respuestas emitidas, R.

Estos estudios no se han podido realizar

pues una cuantificación de la respuesta ante un estímulo no se puede considerar directamente como acertada o no, como en el caso de la proporción de respuestas correctas, PC, obtenidas en los estudios de discriminación sensorial. Sin embargo, se puede desarrollar una medida «indirecta» de PC. El concepto de «indirecta» indica que sería una elaboración sobre las respuestas de estimación de magnitudes y no de una elaboración directa sobre el número de respuestas correctas. Así podríamos definir una medida que indicase la proporción de respuestas ordenadas correctamente. Llamaremos M a esta medida ordinal de respuestas correctas. M se define como el cociente entre el número de no inversiones dadas en las respuestas al orden real de los estímulos (P) partido por el número total de combinaciones de n estímulos tomados de dos en dos ( $C_{n,2}$ ).

$$M = P/C_{n,2} \quad (5)$$

M es útil pues no exige muchos ensayos para su cálculo, se ha utilizado en varias modalidades sensoriales (Garriga, 1985; Garriga y Dorn, 1991; Moreno, 1993; Villarino, 1993) y se han realizado estudios sobre su validación (Garriga, 1987; Villarino, Garriga-Trillo, Rodríguez-Miñón, García-Gallego & Merino, 1991). Así, M se utilizaría como PC ya que es una proporción de respuestas correctas en cuanto a orden se refiere. Con esta medida y la valoración subjetiva media del grado de certeza o confianza, dada explícitamente por cada sujeto, se podría calcular la calibración individual.

Teniendo en cuenta los puntos anteriormente señalados pretendemos en este estudio : (1.º) calcular las calibraciones de los juicios psicofísicos en dos tareas de estimación de magnitudes, (2.º) comparar las tendencias de dichas calibraciones con las obtenidas en tareas «típicamente» cognitivas y con las de tareas sensoriales sólo de discriminación, (3.º) utilizar esta medida para comparar las dos variantes de la técnica de estimación de magnitudes: la estimación de

magnitudes clásica (EMC) y la estimación de magnitudes absoluta (EMA) y (4.º) estudiar las correlaciones entre CA, CO, M, Tiempo medio de reacción (TMR), Edad (E) y variables relacionadas con la función de Stevens, hoy por hoy, la «ley» psicofísica generalmente aceptada.

## Método

### *Sujetos*

En este experimento participaron 30 sujetos, alumnos de Psicología (21 mujeres y 9 hombres), cuyas edades fluctuaban entre los 19 y los 50 años. Su visión era normal y nunca habían participado en experimentos de discriminación visual. Quince sujetos fueron asignados aleatoriamente a cada técnica.

### *Estímulos y Aparatos*

El experimento constaba de siete estímulos que fueron presentados en un ordenador IBM 386 PS2, cuya pantalla medía 12 pulgadas. Las instrucciones fueron dadas de forma escrita y antes de comenzar el experimento. Los siete estímulos consistían en segmentos de recta de las siguientes longitudes en centímetros (cm) : 0,5, 1, 2, 3, 5, 10, 20.

Los estímulos fueron presentados de forma aleatoria en bloques de siete. En cada sesión experimental se presentaban 2 bloques por sujeto presentando en total 420 estímulos. Para emitir las respuestas se utilizaron las teclas numéricas del ordenador. Estas respuestas a su vez pasaban a una base de datos para la realización posterior de los análisis de datos con el SPSS/PC+.

### *Procedimiento*

El experimento se realizó en una sola sesión. En ambas tareas de estimación de magnitudes a cada sujeto se le pedía que juzgase la longitud de las líneas presentadas utilizando números. En la EMC las instrucciones fueron: «En la pantalla del ordenador aparecerán líneas de diferente tamaño en orden

aleatorio. Tu tarea consistirá en juzgar la longitud de las mismas usando números para emitir tus respuestas. Asigna al primer estímulo presentado cualquier número que te parezca apropiado. Para los siguientes estímulos has de utilizar números proporcionales a su longitud aparente, pero teniendo en cuenta el valor asignado al primer estímulo. Por ejemplo, si para el primer estímulo utilizaste el número 10, y la segunda línea te parece el doble de larga, deberías asignarle el número 20; si el segundo estímulo te parece la mitad del primero, deberías asignarle el número cinco y así sucesivamente».

En la EMA, las instrucciones fueron: «En la pantalla del ordenador aparecerán proyectadas líneas de diferente tamaño en orden aleatorio. Tu tarea consistirá en asignar números a cada una de ellas de forma que la magnitud subjetiva del número que elijas, sea igual a la magnitud subjetiva de la longitud de la línea. No intentes estimar la magnitud física de las líneas en centímetros o en cualquier otra unidad de medida. Puedes utilizar cualquier número positivo que te parezca adecuado, es decir, tanto números enteros como decimales o fracciones. Trata cada línea individualmente y no te preocupes por los números que has asignado a las líneas precedentes. Responde tan rápida y espontáneamente como puedas».

Antes de comenzar el experimento se realizaron tantos ensayos de prueba como el sujeto estimase oportuno. Esto se hizo con el fin de ver si se entendían las instrucciones y si se veían bien los estímulos. La sesión duraba entre 15 y 30 minutos.

Todas las sesiones tuvieron lugar en el Laboratorio de Pensamiento y Lenguaje de la Facultad de Psicología de la UNED. Cada sujeto se colocaba frente al ordenador a una distancia de 0.60 metros y leía las instrucciones. Luego, cuando estaba preparado, pulsaba una tecla rotulada «Comienzo». En pantalla aparecía la petición de su nombre, sexo y edad. Dando a la barra espaciadora comenzaban los ensayos de prueba. Des-

pués de cada presentación, y cuando el sujeto lo estimaba oportuno, pulsaba la tecla espaciadora y se le preguntaba primero cuál era su estimación y luego qué confianza tenía en su juicio. La confianza debía representarla en una escala de 0 a 1. Cuando el sujeto lo estimaba oportuno volvía a pulsar la tecla espaciadora y aparecía el segundo estímulo. El procedimiento se repetía hasta el último estímulo.

### Resultados

Las variables consideradas en este trabajo fueron: la confianza media (CO), el coeficiente de determinación obtenido al ajustar los datos (E,R) a la función de Stevens ( $r^2$ ), la ordenada en el origen de dicha función (Inter.), la pendiente (Pend.), el tiempo medio de reacción medido en segundos (TMR), la edad, sexo, la técnica, la medida de la proporción de respuestas ordinales correctas (M) y la calibración (CA). Los valores de todas estas variables en cada sujeto aparecen en la Tabla 1.

Todos los valores obtenidos para CA son negativos, excepto dos que son positivos. Esto también lo podemos observar en las curvas de calibración de la Figura 1. En ella, además, notamos la separación de las curvas de calibración de la EMC y de la EMA de la de calibración perfecta.

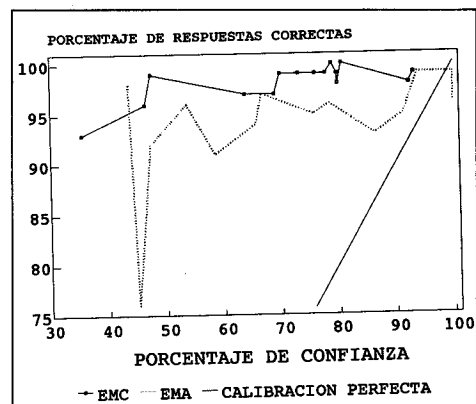


Figura 1: Curvas de calibración.

Tabla 1.  
Valores de las variables estudiadas por sujeto

Sujeto	CO	r <sup>2</sup>	Inter.	Pend.	TMR	Edad	Sexo	Técnica	M	CA
1	.80	.99	-.59	1.02	3.86	19	2	1	1.00	-.20
2	.47	.83	.71	1.32	6.94	19	2	2	.92	-.45
3	.35	.94	1.21	.65	6.19	20	2	1	.93	-.58
4	.59	.80	.45	1.19	3.22	21	1	2	.91	-.32
5	.73	.96	1.41	1.19	7.45	23	2	1	.99	-.26
6	.44	.95	-.48	.91	5.10	25	1	2	.98	-.55
7	.76	.99	1.04	1.11	20.51	20	1	1	.99	-.22
8	.54	.80	.30	.91	1.60	18	2	2	.96	-.42
9	.71	.97	.59	.68	3.59	19	2	1	.99	-.29
10	.79	.78	-.80	1.22	5.25	33	2	2	.96	-.17
11	.64	.91	-.13	.80	4.41	35	2	1	.97	-.33
12	.67	.93	-.16	.99	3.30	35	2	2	.97	-.30
13	.70	.96	-.27	1.07	4.82	23	1	1	.97	-.28
14	.80	.95	.44	.85	2.91	19	2	1	.98	-.18
15	.46	.93	1.10	1.04	6.04	21	2	1	.96	-.49
16	.45	.31	.44	.91	6.94	20	2	2	.76	-.31
17	.78	.98	.68	.99	7.40	34	2	1	.99	-.22
18	1.00	.94	.15	.80	7.63	26	2	2	.96	.04
19	.91	.95	.14	1.04	3.45	29	1	2	.95	-.03
20	.48	.94	-.12	1.38	6.09	32	2	1	.99	-.52
21	.66	.85	.28	.79	7.11	50	2	2	.94	-.28
22	.76	.88	.44	.60	.71	28	1	2	.95	-.19
23	.94	.97	.66	.94	7.92	32	2	2	.99	-.05
24	.93	.99	.05	1.05	2.14	38	1	1	.99	-.06
25	.81	.99	-.47	1.20	5.36	28	2	1	1.00	-.19
26	.86	.84	-.24	.80	5.85	28	2	2	.93	-.07
27	.92	.98	-.16	1.08	4.77	38	1	1	.98	-.06
28	.94	.97	-.22	1.21	5.15	31	2	2	.99	-.05
29	.80	.99	1.71	.96	10.11	29	1	1	.99	-.19
30	1.00	.97	.60	1.05	4.43	31	2	2	.99	.02

Sexo: 1 = Masculino, 2 = Femenino;

Técnica: 1 = Estimación de magnitudes clásica, 2 = Estimación de magnitudes absoluta

Asumiendo que la distribución de las puntuaciones es normal y un nivel de confianza del 95%, los valores a partir de los cuales las calibraciones son significativamente distintas de cero (la CA perfecta) son:  $|CA| \geq .02878$  para la EMC, y para la EMA,  $|CA| \geq .14948$ . Así, a nivel inferencial y en cuanto a técnicas se refiere, la CA

de todos los juicios de EMC son negativos y significativamente distintos de cero. Esto indica que los sujetos manifiestan subconfianza en sus juicios de estimación de magnitudes clásica. En la CA de los juicios de EMA, nueve son significativamente distintos de cero y negativos (manifiestan subconfianza un 60% de la muestra) y seis no

son significativamente distintos de cero, es decir, están bien calibrados. Sin embargo, comparando las calibraciones medias en las dos técnicas obtenemos una  $t(28) = -1$ ,  $p > .05$ , valor que cae fuera de la región crítica eligiendo  $\alpha = .05$ . Por ello, aceptamos que en la población ambas calibraciones medias son iguales. No existen diferencias significativas entre técnicas de estimación de magnitudes en cuanto a calibración se refiere.

Para determinar la relación entre las variables calculamos el coeficiente de correlación de Pearson entre ellas. Los coeficientes calculados sobre datos agrupados por sujeto ( $N = 30$ ) aparecen en la Tabla 2. Estos datos son: la confianza media - CO, la calibración

más baja es la calibración. Como la calibración mejor es la que se acerca a cero, cuando el sujeto tiene más confianza tiende a calibrar mejor y cuando el sujeto tiene menos confianza tiende a calibrar peor. Esta tendencia es coherente con la interpretación de la segunda correlación:  $r_{(CO)(M)}$ . Los sujetos que tienen más confianza en sus juicios tienden a acertar más respuestas, en cuanto a orden se refiere. Este efecto fue encontrado también por Seward (1928) en el reconocimiento de formas. La tercera de las correlaciones relaciona la proporción de respuestas correctas con el grado de ajuste de la función de Stevens. A mayor número de respuestas correctas mejor será el ajuste de Stevens.

Tabla 2  
Matriz de correlaciones (Datos agrupados)

	CO	CA	M	$r^2$	Inter.	Pend.	TMR
CO	1.00	-.97**	.43*	.41	-.17	.03	.01
CA	-.97**	1.00	-.20	-.20	.17	.01	.00
M	.43*	-.20	1.00	.93**	-.08	.15	.03
$r^2$	.41	-.20	.93**	1.00	.04	.08	.06
Inter.	-.17	.17	-.08	.04	1.00	-.16	.42
Pend.	.03	.01	.15	.08	-.16	1.00	.19
TMR	.01	.00	.03	.06	.42	.19	1.00
* ( $p < 0.01$ )				** ( $p < 0.001$ )			

- CA, la proporción de respuestas ordinales correctas - M, el coeficiente de determinación al calcular la función de Stevens -  $r^2$ , la ordenada en el origen de dicha función - Inter., la pendiente de la misma - Pend. y el tiempo medio de reacción medido en segundos - TMR.

Estudiando estas correlaciones sólo obtenemos tres que son significativamente distintas de cero:  $r_{(CO)(CA)}$ ,  $r_{(CO)(M)}$  y  $r_{(M)(G2)}$ . Las variables incluidas en las primeras dos están relacionadas funcionalmente ya que  $CA = CO - M$ . Es curioso notar que la primera de las relaciones es negativa, lo que implica que siendo la relación lineal significativa, a mayor confianza en los juicios

La Tabla 3 presenta las correlaciones calculadas sobre datos sin agrupar: el estímulo - E, la respuesta - R, la confianza asignada a cada juicio - CO, el tiempo de reacción medido en segundos - TR y la edad.

En la matriz de correlaciones de datos no agrupados ( $N = 420$ ) hay cinco correlaciones significativas. La de mayor magnitud indica la relación obvia entre el tamaño del estímulo y la respuesta con la que el sujeto lo cuantifica. El estímulo explica un 40% de la varianza de la respuesta. Las otras cuatro, aunque significativas, explican un porcentaje muy bajo de la varianza total en cada comparación y siguiendo las indicaciones de Cohen (1990) no las comentamos debido a

Tabla 3  
Matriz de correlaciones (Datos no agrupados)

	E	R	CO	TR	EDAD
E	1.00	-.63**	.02	.15**	.00
R	.63**	1.00	-.09	.20**	-.14*
CO	.02	-.09	1.00	-.09	.26**
TR	.15**	.20**	-.09	1.00	-.05
EDAD	.00	-.14*	.26**	-.05	1.00
		* (p<0.01)		** (p<0.001)	

que su significación es función del tamaño de la muestra siendo la magnitud del efecto en sí muy bajo.

### Discusión

El efecto de subconfianza observado en nuestro estudio de la calibración de juicios psicofísicos de estimación de magnitudes es bastante estable. Se da en todos, menos dos, de los sujetos y en las dos técnicas aplicadas, no presentando diferencias significativas entre los valores medios obtenidos para cada técnica.

Además, tanto la confianza como la calibración, representan variables «nuevas» en este tipo de escalamiento en el sentido de que miden aspectos no considerados en las variables habituales en estas tareas como son la pendiente, la ordenada y el grado de ajuste de la función psicofísica de Stevens y los tiempos de reacción. La misma proporción de respuestas correctas utilizada no se relaciona significativamente con la calibración. Esta aparente disyunción entre los aspectos tradicionales del escalamiento sensorial y estos aspectos «más» cognitivos podrían revelar que los procesos subyacentes a la emisión de juicios sean distintos según la tarea a realizar. Aunque el cuantificar un estímulo requiere un juicio sobre una base de conocimientos incompleta, al igual que en el juicio sobre la confianza en la respuesta, ambos procesos podrían ser distintos y por eso ambas respuestas no correlacionan.

Un proceso indica un juicio directo (la respuesta ante un estímulo) y el otro sería un metajuicio (un juicio sobre la certeza del juicio directo). Esta diferencia entre procesos explicaría también la diferencia entre el efecto de sobreconfianza establemente encontrado como ilusión cognitiva (Gigerenzer, Hoffrage y Kleinbölting, 1991) y el efecto de subconfianza hallado en este estudio y en los relacionados con la discriminación sensorial de Björkman et al (1993). Si aceptásemos la división de Howell (1971) entre dos tipos de incertidumbre, podríamos sugerir que la sobreconfianza surge de la incertidumbre interna y la subconfianza de la incertidumbre externa. Así, en el caso de los juicios «más» cognitivos, tales como las tareas de conocimiento general o los juicios expertos, la incertidumbre sería inherente al evaluador. En los juicios psicofísicos, la incertidumbre se encontraría en el mundo externo, donde se encuentra aquello sobre lo que se tiene que emitir el juicio.

En base a nuestros resultados, la medida M nos conduce hacia conclusiones coherentes con las investigaciones realizadas sobre la calibración de los juicios psicofísicos de discriminación y con las que determinan su relación positiva con el aspecto de la confianza en los juicios emitidos en cualquier tipo de tarea. Por ello juzgamos que es adecuada, además de como medida ordinal de la sensibilidad, para estos estudios de calibración en juicios psicofísicos de magnitudes.

Referencias

- Björkman, M., Juslin, P. & Winman, A. (1993). Realism of confidence in sensory discrimination: The underconfidence phenomenon. *Perception & Psychophysics*, 54, 75-81.
- Cohen, J. (1990). Things I have learned (So far). *American Psychologist*, 45, 1304-1312.
- Cooper, A. C, Woo, C. Y. & Dunkelberg, W. C. (1988). Entrepreneurs' perceived chances of success. *Journal of Business Venturing*, 3, 97-108.
- Garriga Trillo, A. (1985). *Función Psicofísica y Medida de la Sensibilidad Olfativa*. Tesis Doctoral no publicada, Universidad Autónoma de Madrid, Madrid. Resúmenes de la UNESCO 1985.
- Garriga Trillo, A. (1987). Olfactory psychophysics: Sensitivity measures. En E. E. Roskam y R. Suck (Eds.) *Progress in Mathematical Psychology-1* (pp 343-349). Amsterdam: North-Holland.
- Garriga Trillo, A. J. y Dorn, T. (1991). Medición de la borrosidad: Modalidades cruzadas. *Psicothema*, 32, 423-432.
- Gigerenzer, G. Hoffrage, U. y Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, 98, 4, 506-528
- Glaser, R. (1989). Expertise and Learning: How do we think about instructional processes now that we have discovered knowledge structures? En D. Klahr y K. Kotovsky (Eds.) *Complex Information Processing: The Impact of Herbert A. Simon*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Howell, W. C. (1971). Uncertainty from internal and external sources: A clear case of overconfidence. *Journal of Experimental Psychology*, 89, 240-243.
- Lichtenstein, S. & Fischhoff, B. (1977). Do those who know more also know more about how much they know? The calibration of probability judgments. *Organizational Behavior and Human Performance*, 20, 159-183.
- Lichtenstein, S., Fischhoff, B., & Phillips, L.D. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment Under Uncertainty: Heuristics and Biases* (pp 306-334). Cambridge: Cambridge University Press.
- Lusted, L. B. (1977). *A Study of the Efficacy of Diagnostic Radiologic Procedures: Final Report on Diagnostic Efficacy*. Chicago: Efficacy Study Committee of the American College of Radiology.
- Moreno, E. (1993). [Medidas de la sensibilidad en la modalidad visual]. Datos no publicados.
- Oskamp, S. (1965). Overconfidence in case-study judgments. *The Journal of Consulting Psychology*, 29, 261-265.
- Seward, G. (1928). Recognition time as a measure of confidence. *Archives of Psychology*, 99.
- Skotnikova, I. G. (1993, October). *Confidence-Unconfidence and reaction time in unordered discrimination: Subject oriented research*. Paper presented at the Ninth Meeting of the International Society for Psychophysics, Palma de Mallorca, Spain.
- Villarino, A. (1993). [Medidas de la sensibilidad en la modalidad gustativa]. Datos no publicados.
- Villarino, A., Garriga-Trillo, A., Rodríguez-Miñón, P., García-Gallego, C. & Merino, J. M. (1991). Comparing sensitivity indices using different models and modalities. En G. Lockhead *Fechner Day 91* (pp 95-98). Durham, North Carolina: International Society for Psychophysics.
- Yates, J. F. (1990). *Judgment and Decision Making*. Englewoods Cliffs, NJ: Prentice Hall.
- Wagenaar, W. A. & Keren, G. (1986). Does the expert know? The reliability of predictions and confidence ratings of experts. In E. Hollnagel, G. Maneini, & D. Woods (Eds.), *Intelligent Decision Support in Process Environments* (pp. 87-107). Berlin: Springer.

Aceptado el 7 de febrero. de 1994