

Disolución del sesgo de subconfianza en tareas verbales

Guillermo Macbeth y Eugenia Razumiejczyk*

Instituto de Investigaciones Psicológicas de la Universidad del Salvador (IIPUS) (Buenos Aires, Argentina)

Resumen: Se define el sesgo de subconfianza como la subestimación del éxito subjetivo en comparación con el éxito objetivo de un agente en una serie de tareas. Se presenta un experimento que disuelve el sesgo de subconfianza en tareas verbales mediante dos intervenciones cognitivas sencillas. La primera intervención se propone entrenar a los participantes en tareas verbales, la segunda se propone entrenarlos en calibración sobre tareas verbales. La muestra se conformó con 210 estudiantes universitarios. Se pusieron a prueba cinco hipótesis vinculadas a la disolución del sesgo de subconfianza. Se conformaron aleatoriamente dos grupos experimentales y un grupo control. En fase de estudio se entrenó a un grupo experimental en tareas verbales y al otro en calibración sobre tareas verbales. En fase de prueba se administró un test de aptitud verbal a los tres grupos, seguido de una tarea de éxito subjetivo. Los resultados indican que ambas intervenciones disuelven el sesgo de subconfianza, pero la segunda es más eficiente que la primera. Se sugiere, adicionalmente, que el entrenamiento en tareas específicas disuelve el sesgo cuando ofrece la ocasión de entrenar indirectamente la calibración.

Palabras clave: Calibración; disolución; sesgo de subconfianza; tareas verbales.

Title: Debiasing overconfidence bias in verbal tasks.

Abstract: The underconfidence bias is defined as the underestimation of the subjective success in comparison with the objective or actual success achieved by an agent in a series of tasks. An experiment that dissolves the underconfidence bias in verbal tasks by means of two simple cognitive interventions is presented. The aim of the first debiasing intervention is to train the participants in verbal tasks; the aim of the second is to train them in calibration on verbal tasks. 210 university students participated in the experiment. Five experimental hypotheses related to the underconfidence bias debiasing were tested. Subjects were randomly assigned to a control group or to one of two experimental groups. In study phase one experimental group was trained in verbal tasks and the other in calibration on verbal tasks. In test phase a verbal aptitude test was administered to the three groups, followed by a subjective estimation of success task. Results show that the second cognitive intervention is more efficient than the first to achieve the debiasing of the underconfidence bias. It is suggested, in addition, that the training in specific tasks dissolves the calibration bias when it offers the occasion to train indirectly the cognitive processes of calibration.

Key words: Calibration; debiasing; underconfidence bias; verbal tasks.

1. Introducción

Se definen los sesgos de sobreconfianza y subconfianza como distorsiones de la calibración o confianza (Camerer y Lovallo, 1999; Finn y Metcalfe, 2007; Lichtenstein, Fischhoff y Phillips, 1982; Oskamp, 1965; Simmons y Nelson, 2006). Formalmente, la calibración o confianza puede ser definida como la relación entre el éxito subjetivo y el éxito objetivo que presenta la Ecuación 1.

$$C = \sum_{i=1}^n (E_i - O_i) \quad (1)$$

La calibración o confianza C se obtiene de restar el éxito real u objetivo O_i del éxito subjetivo o estimado E_i de un agente en una serie de tareas específicas i . La sobreconfianza se observa cuando la suma que define a C genera puntuaciones positivas, por lo que el éxito subjetivo resulta mayor que el éxito objetivo ($E > O$). La subconfianza, por el contrario, se observa cuando la C resulta negativa, por lo que el éxito subjetivo es menor que el objetivo ($E < O$). El agente se encuentra bien calibrado cuando la diferencia entre E_i y O_i resulta igual a cero. En tal caso, la C no es positiva ni es negativa, es decir, no se observan sesgos.

Los sesgos de sobreconfianza y subconfianza han sido considerados como distorsiones típicas del procesamiento cognitivo de las decisiones humanas bajo incertidumbre (Kahneman y Tversky, 1996; Koheler, Brenner y Griffin, 2002; Merkle y Van Zandt, 2006). Sin embargo, sostienen

algunos autores que la ocurrencia de estos sesgos puede ser considerada como un fenómeno de laboratorio que no llega a generar desajustes significativos en la adaptación ecológica de los individuos a su medio (Dhimi, Hertwig y Hoffrage, 2004; Gigerenzer, Hoffrage y Kleinböling, 1991; Klayman, Soll, Gonzalez-Vallejo y Barlas, 1999). Las distorsiones de la calibración se han observado, no obstante, en diversas poblaciones y en múltiples tareas (Garriga Trillo, Villarino, González Labra y Arnau, 1994; Koheler *et al.*, 2002; Lichtenstein *et al.*, 1982; Macbeth, 2005, 2006; Macbeth y Cortada de Kohan, en prensa; Macbeth, Cortada de Kohan, Razumiejczyk y López Alonso, 2006; Macbeth, Razumiejczyk y Cortada de Kohan, 2006; Merkle y Van Zandt, 2006).

El sesgo de sobreconfianza ha sido ampliamente estudiado en relación con la economía (Angner, 2006; Camerer y Lovallo, 1999; Macbeth y Cortada de Kohan, en prensa), el diagnóstico médico (Christensen-Szalanski y Bushyhead, 1981; Gigerenzer *et al.*, 1991), el diagnóstico psicológico (Oskamp, 1965), el pronóstico meteorológico (Murphy y Winkler, 1974, 1977), la estimación de magnitudes físicas (Henrion y Fischhoff, 1986) y psicofísicas (Garriga Trillo *et al.*, 1994), los conocimientos generales (Lichtenstein *et al.*, 1982; Macbeth *et al.*, 2006), los trastornos de ansiedad (Macbeth y Bogiaizian, 2007; Wilson, Centerbar y Brekke, 2002), la aptitud matemática (Macbeth, Cortada de Kohan y Kohan Cortada, 2007) y la aptitud verbal (Macbeth, *et al.*, 2006), entre otros ámbitos (Clayson, 2005; Koheler *et al.*, 2002; Svenson, 1981).

El sesgo de subconfianza, en cambio, ha sido poco estudiado. Los experimentos clásicos de Oskamp (1965) y Lichtenstein *et al.* (1982), un estudio reciente de Finn y Metcalfe (2007) y las revisiones de Fischhoff (1982) y Koheler *et al.* (2002) sugieren que su ocurrencia se relaciona con la facili-

* Dirección para correspondencia [Correspondence address]: Guillermo Macbeth. Instituto de Investigaciones Psicológicas de la Universidad del Salvador (IIPUS). Marcelo T. de Alvear 1314 (C1058AAV). Ciudad Autónoma de Buenos Aires (Argentina).
E-mail: guillermo.macbeth@mail.salvador.edu.ar

dad de las tareas como variable independiente crítica. Se ha encontrado que el sesgo de subconfianza se genera cuando las tareas experimentales sobre las que se realiza la calibración resultan relativamente fáciles. Estudios recientes han replicado estos hallazgos en tareas de conocimientos generales (Macbeth *et al.*, 2006).

La disolución de los sesgos de la calibración se ha intentado, o bien mediante estrategias metodológicas (Brenner, 2000, 2003; Erev, Wallsten y Budescu, 1994; Klayman *et al.*, 1999; Wallsten, Erev y Budescu, 2000), o bien mediante intervenciones cognitivas (Dhami *et al.*, 2004; Fischhoff, 1982; Gigerenzer *et al.*, 1991; Kahneman y Tversky, 1982, 1996; Koheler *et al.*, 2002; Macbeth, 2006; Macbeth y Bogiazian, 2007; Macbeth *et al.*, 2007).

La revisión del estado de la cuestión sugiere que el sesgo de subconfianza resulta, en términos generales, un fenómeno robusto (Kahneman, 2003) y difícil de disolver (Fischhoff, 1982; Kahneman y Tversky, 1982; Koheler *et al.*, 2002; Liberman, 2004; Liberman y Tversky, 1993). No obstante, Gigerenzer *et al.* (1991) y Dhami *et al.* (2004) sostienen que tanto el sesgo de sobreconfianza, como el sesgo de subconfianza pueden disolverse fácilmente al presentar las tareas experimentales en un formato más ecológico que artificial o de laboratorio, es decir, bajo una modalidad que replique la interacción habitual de los participantes con su entorno cotidiano. Esta estrategia de disolución del sesgo de subconfianza ha resultado parcialmente exitosa (Dhami *et al.*, 2004).

El propósito de este estudio es lograr la disolución del sesgo de subconfianza desde una perspectiva diferente, aunque complementaria. Se pretende eliminar este sesgo mediante intervenciones cognitivas vinculadas a entrenamientos experimentales simples que puedan aplicarse dentro y también fuera del laboratorio (Koheler *et al.*, 2002). Se considera que la disolución de los sesgos de la calibración depende de procesos cognitivos y metacognitivos específicos que pueden optimizarse mediante el entrenamiento (Kahneman, 2003). Se postula que tanto el entrenamiento en la tarea específica, como el entrenamiento focalizado directamente en la calibración, disuelven el sesgo de subconfianza en tareas verbales. Se postula, asimismo, que el entrenamiento en calibración es más eficiente que el entrenamiento en la tarea verbal específica para la disolución del sesgo de subconfianza.

2. Método

2.1. Participantes

Participaron en el estudio 210 estudiantes argentinos reclutados de cursos universitarios de grado y postgrado. La edad promedio de los participantes resultó de 22.34 años ($de = 5.769$ años). La población femenina resultó de 107 mujeres (51%) y la masculina de 103 varones (49%). Para realizar las mediciones se implementaron seis sesiones grupales ($30 < n < 40$) en diferentes condiciones de manipulación experimental.

2.2. Instrumento

Se aplicó a todos los participantes el Test Baires (Cortada de Kohan, 2003) de aptitud verbal. El Test Baires es una prueba de papel y lápiz que presenta tareas de sinónimos y definiciones. En cada ítem se ofrecen cuatro opciones de respuesta, de las cuales una y sólo una es la correcta. La prueba se construyó mediante las técnicas de la Teoría de Respuesta al Ítem (Barbero García, Prieto Marañón, Suárez Falcón y San Luis Costas, 2001; Cortada de Kohan, 1999; Hambleton, 1994) y posee adecuados niveles de confiabilidad y validez. El Test Baires presenta dos versiones que poseen propiedades psicométricas similares, una extensa y otra breve. La versión extensa cuenta con 98 ítems. La versión breve, empleada en este estudio, cuenta con 34 ítems, de los cuales 17 presentan la tarea de encontrar definiciones y los 17 restantes presentan la tarea de encontrar sinónimos. Se empleó, adicionalmente, a modo de entrenamiento experimental en la misma tarea verbal, una versión reducida compuesta por 16 ítems extraídos al azar de la versión extensa, de los cuales 8 presentan tareas con definiciones y 8 con sinónimos. La versión breve y la versión reducida no comparten ningún ítem. En ambos casos, el coeficiente α de Cronbach resultó mayor a 0.70. Las dos versiones del Test Baires empleadas en este estudio, es decir la breve y la reducida, han resultado adecuadas para la generación experimental del sesgo de subconfianza en estudios previos (Macbeth, 2006; Macbeth *et al.*, 2006).

2.3. Procedimiento

Se determinaron tres grupos aleatorios de participantes para conformar un grupo control y dos grupos experimentales. En fase de estudio, los dos grupos experimentales recibieron distinto tratamiento y el grupo control no participó. En fase de prueba, los tres grupos recibieron la misma tarea y consigna sobre la versión breve del Test Baires.

El grupo control ($n = 77$) no recibió ninguna manipulación en fase de estudio. En fase de prueba los participantes recibieron la versión breve del Test Baires y una consigna de éxito subjetivo, para la cual debieron indicar cuántos de los 34 ítems creen haber respondido correctamente. La calibración C se operacionalizó como la discrepancia entre el éxito subjetivo y el éxito objetivo indicada en la Ecuación 1.

El grupo experimental 1 ($n_{EXPI} = 61$) recibió, en fase de estudio, un entrenamiento en la tarea verbal específica implementado mediante la administración de una forma reducida del Test Baires de 16 ítems. La consigna del grupo experimental 1 consistió sólo en completar la tarea verbal. El experimentador señaló, al finalizar la fase de estudio, las respuestas correctas al Test Baires empleado. El objetivo de la intervención aplicada al grupo experimental 1 fue entrenar a los participantes en la tarea verbal específica. La fase de prueba del grupo experimental 1 fue idéntica a la del grupo control.

El grupo experimental 2 ($n_{EXP2} = 72$) recibió, en fase de estudio, un entrenamiento en calibración sobre la base de la misma forma reducida del Test Baires que se empleó en el grupo experimental 1. Se solicitó a los participantes, en primer lugar, que completen la prueba verbal y, luego, que realicen una tarea de estimación de éxito subjetivo. Esta tarea consistió en estimar cuántos de los 16 ítems fueron respondidos correctamente. Luego de completar esta tarea de éxito subjetivo, el evaluador señaló las respuestas correctas y consignó a los participantes que consideren brevemente las discrepancias obtenidas entre el éxito objetivo y el éxito subjetivo a modo de entrenamiento en calibración. La manipulación implementada en el grupo experimental 2 se propuso mejorar el rendimiento en calibración, es decir, promover una máxima proximidad de C a cero. La fase de prueba del grupo experimental 2 fue idéntica a la del grupo control.

2.4. Hipótesis

Se pusieron a prueba cinco hipótesis vinculadas a la disolución del sesgo de subconfianza en tareas verbales.

La hipótesis H_1 afirma la ocurrencia del sesgo de subconfianza en los participantes que no reciben manipulación experimental, es decir, en el grupo control. Formalmente, $H_1 : \bar{E}_{CONTROL} < \bar{O}_{CONTROL}$. Este predominio de la media del éxito objetivo O por sobre la media del éxito subjetivo E se pronostica para tareas relativamente fáciles, en coherencia con estudios previos (Gigerenzer *et al.*, 1991; Lichtenstein *et al.*, 1982; Macbeth *et al.*, 2006). Se esperan puntuaciones de calibración C menores que cero por el predominio del éxito objetivo O por sobre el éxito subjetivo E cuando los participantes no reciben ningún entrenamiento.

La hipótesis H_2 sostiene que el entrenamiento en la tarea verbal sobre la que se mide el éxito objetivo O disuelve el sesgo de subconfianza. Se pronostica una diferencia no significativa entre la media del éxito objetivo O y la media del éxito subjetivo E para los participantes entrenados en O , es decir, para el grupo experimental 1 (EXP1). Formalmente, $H_2 : \bar{E}_{EXP1} \approx \bar{O}_{EXP1}$. Se esperan puntuaciones de calibración C próximas a cero por la manipulación implementada en O . La disolución del sesgo de subconfianza por entrenamiento en la tarea verbal sobre la que se mide el éxito objetivo se justifica por el ajuste indirecto de la calibración que se genera mediante la práctica en la misma tarea verbal específica. La hipótesis H_2 es coherente con hallazgos anteriores (Fischhoff, 1982; Macbeth y Bogaizian, 2007; Macbeth *et al.*, 2007; Murphy y Winkler, 1974, 1977).

La hipótesis H_3 afirma que el entrenamiento en calibración C disuelve el sesgo de subconfianza. La hipótesis H_3 pronostica una diferencia no significativa entre la media del éxito objetivo O y la media del éxito subjetivo E para los participantes entrenados en C , es decir, para el grupo experimental 2 (EXP2). Formalmente, $H_3 : \bar{E}_{EXP2} \approx \bar{O}_{EXP2}$. Se esperan puntuaciones de calibración C próximas a cero

por la manipulación experimental orientada a la optimización de los procesos cognitivos y metacognitivos que regulan la calibración (Kahneman, 2003). La hipótesis H_3 es coherente con estudios previos (Lichtenstein *et al.*, 1982; Macbeth *et al.*, 2007) y se justifica por el ajuste progresivo de la calibración que se genera por el entrenamiento focalizado directamente en C .

La hipótesis H_4 afirma que la disolución del sesgo de subconfianza que se logra mediante entrenamiento en calibración C es más eficiente que la disolución que se logra mediante entrenamiento en la tarea verbal sobre la que se mide el éxito objetivo O . La hipótesis H_4 pronostica que los puntuaciones de C se aproximan más a cero por efecto del entrenamiento focalizado en C del grupo experimental 2 que por efecto del entrenamiento focalizado en O del grupo experimental 1. Formalmente,

$H_4 : \bar{C}_{CONTROL} > \bar{C}_{EXP1} > \bar{C}_{EXP2}$. La hipótesis H_4 es coherente con hallazgos anteriores (Fischhoff, 1982; Gigerenzer *et al.*, 1991, Macbeth *et al.*, 2006) y se justifica por la diferencia funcional que sugieren estudios previos entre los procesos cognitivos responsables del éxito objetivo O y los responsables de la calibración C .

La hipótesis H_5 sostiene que la magnitud de los sesgos de la calibración es mínima cuando los participantes han recibido entrenamiento en calibración C , máxima cuando no han recibido ningún tipo de entrenamiento, e intermedia cuando han recibido entrenamiento en la tarea verbal específica sobre la que se mide el éxito objetivo O . Así, la hipótesis H_5 pronostica que el tamaño del efecto (TE) es mayor cuando los participantes no han recibido ningún entrenamiento, menor cuando han recibido entrenamiento en O , y menor aún cuando han recibido entrenamiento en C . Formalmente, $H_5 : TE_{CONTROL} > TE_{EXP1} > TE_{EXP2}$. Esta hipótesis se justifica por la diferencia funcional que sugieren los resultados de estudios previos entre los procesos cognitivos responsables del éxito objetivo O y los responsables de la calibración C . La hipótesis H_5 es una variante de la hipótesis H_4 que permite estudiar de manera complementaria la disolución del sesgo de subconfianza propuesta en este estudio.

3. Resultados y discusión

La hipótesis H_1 resultó coherente con la evidencia experimental. El éxito subjetivo ($\bar{E}_{CONTROL} = 17.10$; $d.t. = 5.248$) resultó significativamente menor que el éxito objetivo ($\bar{O}_{CONTROL} = 19.84$; $d.t. = 3.873$) en el grupo control por la prueba t para muestras apareadas ($t = 7.0$; $p < .001$; $d = 0.594$). Los participantes que no recibieron ninguna manipulación incurrieron en el sesgo de subconfianza. Este hallazgo es coherente con estudios previos que registraron la subestimación del éxito subjetivo en comparación con el éxito objetivo en tareas relativamente fáciles de conocimientos generales (Lichtenstein *et al.*, 1982; Macbeth *et al.*, 2006).

Hallazgos anteriores sugieren que las tareas relativamente difíciles tienden a generar el sesgo de sobreconfianza (Angner, 2006; Camerer y Lovallo, 1999; Cortada de Kohan y Macbeth, 2006; Lichtenstein *et al.*, 1982) mientras que las tareas relativamente fáciles tienden a generar el sesgo de subconfianza (Erev *et al.*, 1994; Juslin, Winman y Olsson, 2000; Koheler *et al.*, 2002; Oskamp, 1965).

La hipótesis H_2 resultó coherente con la evidencia experimental. La discrepancia entre el éxito subjetivo ($\bar{E}_{EXP1} = 20.11$; $d.t. = 3.971$) y el éxito objetivo ($\bar{O}_{EXP1} = 19.30$; $d.t. = 2.629$) en el grupo experimental 1 resultó no significativa por la prueba t para muestras apareadas ($t = -1.588$; $p = .118$; $d = 0.241$). Los participantes que en fase de estudio recibieron un entrenamiento simple en la tarea verbal sobre la que se midió el éxito objetivo O no incurrieron, en fase de prueba, en el sesgo de subconfianza. La subestimación del éxito subjetivo en comparación con el éxito objetivo que se observó en el grupo control logró disolverse en el grupo experimental 1 mediante una manipulación cognitiva focalizada en O . Si bien el entrenamiento en O resultó eficaz para disolver el sesgo de subconfianza, el éxito objetivo no mejoró realmente con el entrenamiento implementado en el grupo experimental 1. La diferencia entre la media de O en el grupo experimental 1 y la media de O en el grupo control resultó no significativa por la prueba t para muestras independientes ($t = 0.947$; $p = .345$; $d = 0.16$). Este hallazgo es novedoso en tanto aporta evidencia acerca de la posibilidad de disolución del sesgo de subconfianza mediante un entrenamiento simple en O , aún cuando la intervención no logre mejorar el éxito objetivo. Estudios previos, no obstante, señalan la resistencia de los sesgos de la calibración a estrategias de disolución vinculadas a la advertencia directa a los participantes (Fischhoff, 1982), al entrenamiento en probabilidad (Lichtenstein *et al.*, 1982) y a la adquisición de experticia (Angner, 2006; Oskamp, 1965). Se considera en este estudio que el entrenamiento en O implementado en el grupo experimental 1, si bien no incluye consignas que activen directamente los procesos de éxito subjetivo E , ofrece la ocasión de entrenar de manera indirecta la calibración C . Se sugiere que éste es el motivo por el cual se disolvió el sesgo de subconfianza en el grupo experimental 1. Aunque el entrenamiento en tareas verbales no mejoró el éxito objetivo, la calibración dejó de presentar sesgos. De esta manera, la facilidad de la tarea se mantuvo estable, pero el sesgo se disolvió. Así, la facilidad relativa de la tarea no puede ser considerada, en el presente estudio, como una variable independiente crítica del sesgo de subconfianza.

La hipótesis H_3 resultó coherente con la evidencia experimental. La discrepancia entre el éxito subjetivo ($\bar{E}_{EXP2} = 21.82$; $d.t. = 6.211$) y el éxito objetivo ($\bar{O}_{EXP2} = 22.44$; $d.t. = 4.497$) en el grupo experimental 2 resultó no significativa por la prueba t para muestras apareadas ($t = -1.009$; $p = .317$; $d = 0.114$). Los participantes que en fase de estudio recibieron entrenamiento en calibración sobre tareas verbales no incurrieron, en fase de prueba, en el sesgo de subconfianza. La

subestimación del éxito subjetivo en comparación con el éxito objetivo que se observó en el grupo control logró disolverse mediante una manipulación experimental focalizada en la aproximación de C a cero. Este resultado es coherente con estudios previos que reportan la ausencia de sesgos de la calibración en meteorólogos que hacen pronósticos diariamente y conocen luego, a corto plazo, su acierto o desacierto (Fischhoff, 1982; Murphy y Winkler, 1974, 1977). Esta actividad permite, posiblemente, entrenar los procesos cognitivos y metacognitivos del éxito subjetivo que regulan la calibración. La disolución del sesgo de subconfianza en tareas verbales mediante entrenamiento en calibración lograda en el presente estudio es, no obstante, original.

Si bien el objetivo específico de la manipulación implementada en el grupo experimental 2 fue la optimización de la calibración C , es decir, la promoción de su proximidad a cero, también el éxito objetivo O resultó críticamente afectado. El grupo experimental 2 obtuvo una media de O significativamente superior al grupo control por la prueba t para muestras independientes ($t = -3.789$; $p < .001$; $d = 0.621$). Este resultado sugiere, adicionalmente, que el entrenamiento en calibración sobre tareas verbales es más eficiente que el entrenamiento en tareas verbales para la optimización del éxito objetivo en esas tareas específicas.

La hipótesis H_4 resultó coherente con la evidencia experimental. La calibración C , definida según la Ecuación 1, resultó más discrepante de cero en el grupo control ($\bar{C}_{CONTROL} = -2.74$; $d.t. = 3.435$) que en el grupo experimental 1 ($\bar{C}_{EXP1} = 0.82$; $d.t. = 4.031$). La calibración observada en el grupo experimental 2 ($\bar{C}_{EXP2} = -0.63$; $d.t. = 5.258$) resultó, a la vez, más próxima a cero que la calibración observada en el grupo experimental 1. Las diferencias de calibración entre los tres grupos resultaron significativas según el ANOVA de una vía [$F(2,207) = 12.017$; $p < .001$; $\eta^2 = .104$]. En coherencia con la hipótesis H_4 , la disolución del sesgo de subconfianza resultó más eficiente por entrenamiento en C que por entrenamiento en O . La proximidad de la calibración a cero es mayor en el grupo experimental 2 que en el grupo experimental 1. El comportamiento de la calibración C sin manipulación (grupo control), con entrenamiento en O (grupo experimental 1) y con entrenamiento en C (grupo experimental 2) puede apreciarse gráficamente en la Figura 1.

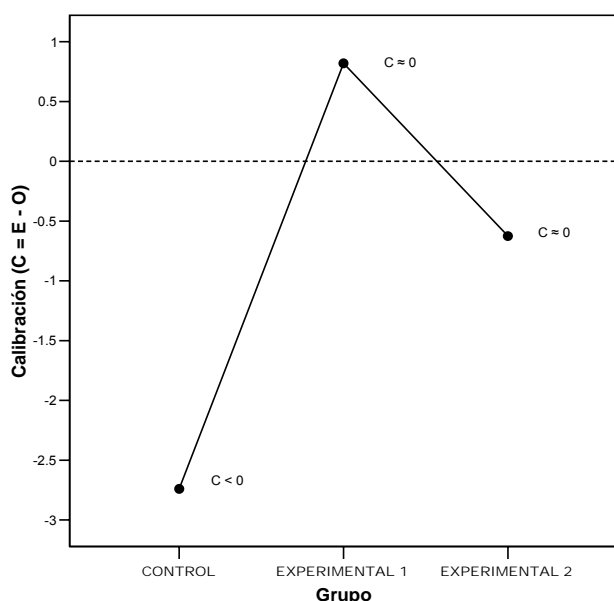


Figura 1: El comportamiento de la calibración en los grupos control, experimental 1 y experimental 2

Si bien las puntuaciones de la calibración C se volvieron positivas ($E > O$) por efecto del entrenamiento en tareas verbales (grupo experimental 1) y se mantuvieron negativas ($E < O$) por efecto del entrenamiento en calibración (grupo experimental 2), la diferencia entre el éxito subjetivo E y el éxito objetivo O resultó no significativa en ambos casos. Sólo el grupo control presentó el sesgo de subconfianza. Los participantes de los grupos experimentales 1 y 2, en cambio, no incurrieron en ningún sesgo de la calibración. En la Figura 1 se aprecia la proximidad a cero de los grupos experimentales 1 y 2 y la distancia a cero del grupo control.

Este resultado es original y promueve la distinción funcional entre los procesos cognitivos responsables del éxito objetivo en tareas verbales y los procesos cognitivos responsables de la calibración. Si bien logró disolverse el sesgo en el grupo experimental 1, se sugiere en este estudio que la experticia logra disolver los sesgos de la calibración sólo en ciertas ocasiones porque su adquisición interviene de manera indirecta sobre los procesos cognitivos y metacognitivos de la calibración (Kahneman, 2003).

La hipótesis H_3 resultó coherente con la evidencia experimental. El tamaño del efecto (TE) de la calibración C resultó mayor para el grupo control, menor para el grupo experimental 2, e intermedia para el grupo experimental 1. De esta manera, la calibración presenta la magnitud máxima de sesgo en el grupo que no recibió ningún entrenamiento, la magnitud mínima en el grupo que recibió entrenamiento en calibración y una magnitud intermedia en el grupo que recibió entrenamiento en tareas verbales.

Se calculó el TE correspondiente a los grupos control, experimental 1 y experimental 2 mediante la d de Cohen (1988), la g de Hedges (Grissom y Kim, 2005), la Δ de Glass

(Glass, McGaw y Smith, 1981), el estadístico CL o *common language effect size* (McGraw y Wong, 1992; Valera Espín y Sánchez Meca, 1997) y la conversión de d a r propuesta por Cohen (1988). La Tabla 1 resume el TE para la diferencia C entre el éxito subjetivo E y el éxito objetivo O correspondiente a los tres grupos. En todos los casos se observa que el TE de la calibración es mayor en el grupo control, menor en el grupo experimental 2, e intermedia en el grupo experimental 1.

Tabla 1: Medidas del tamaño del efecto para la calibración en los tres grupos.

TE Calibración	Grupo		
	Control	Experimental 1	Experimental 2
d	0.594	0.241	0.114
g	0.594	0.241	0.114
Δ	0.522	0.204	0.100
CL	0.663	0.568	0.532
d a r	0.285	0.119	0.057

La comparación gráfica de los tres grupos según la diferencia estandarizada entre la media de E y la media de O para la calibración C que provee la d de Cohen (1988) se presenta en la Figura 2. Se observa que la magnitud del sesgo de subconfianza es máxima en el grupo control y que su disolución es más eficiente mediante el entrenamiento en calibración implementado en el grupo experimental 2 que mediante el entrenamiento en la tarea verbal implementado en el grupo experimental 1.

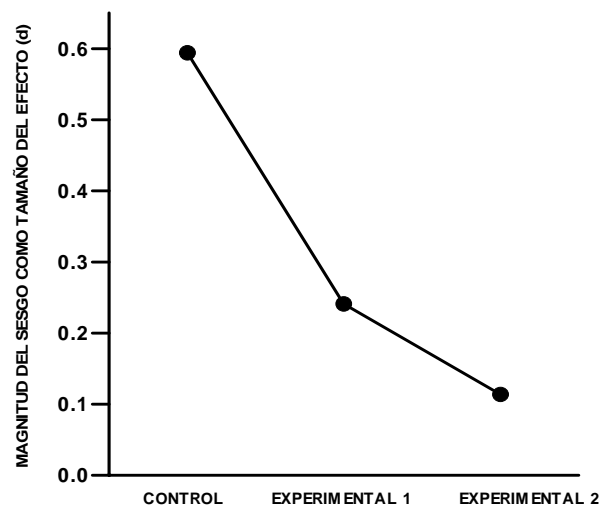


Figura 2: La disolución del sesgo de subconfianza como reducción del tamaño del efecto

Las distribuciones de frecuencias de las variables comparadas en las sucesivas pruebas de hipótesis resultaron compatibles con la presunción de normalidad y homocedasticidad por las pruebas de Kolmogorov-Smirnov y Levene, respectivamente.

4. Conclusiones

El sesgo de subconfianza en tareas verbales que se observó espontáneamente en ausencia de manipulación logró disolverse mediante dos formas sencillas de intervención cognitiva, una centrada en tareas verbales, otra centrada en calibración sobre las mismas tareas específicas. El entrenamiento en calibración resultó más eficiente que el entrenamiento restringido a las tareas verbales para la disolución del sesgo de subconfianza. Adicionalmente, el éxito objetivo en tareas verbales no logró mejorar con un entrenamiento focalizado en tareas verbales similares, pero se incrementó de manera significativa con un entrenamiento focalizado en calibración.

Estos hallazgos promueven la distinción funcional entre los procesos cognitivos responsables del rendimiento en tareas específicas y los responsables del rendimiento en calibración. El aporte específico de estos hallazgos es original y provee evidencia relevante para la comprensión del sesgo de subconfianza y su disolución mediante estrategias cognitivas sencillas. Al tratarse de intervenciones simples, se considera que pueden aplicarse tanto dentro (Merkle y Van Zandt,

2006), como fuera del laboratorio (Koheler *et al.*, 2002), por ejemplo, en tareas económicas (Angner, 2006; Macbeth y Cortada de Kohan, en prensa), educativas (Clayson, 2005) y clínicas (Christensen-Szalanski y Bushyhead, 1981; Lichtenstein *et al.*, 1982; Oskamp, 1965), entre otras (Svenson, 1981).

En síntesis, se concluye que la disolución del sesgo de subconfianza que se observa en tareas verbales relativamente fáciles es más eficiente cuando se implementa un entrenamiento en calibración que cuando se implementa un entrenamiento en las mismas tareas verbales específicas. Se considera que este fenómeno se debe a que la retroalimentación de los resultados que implementa el entrenamiento en calibración permite un ajuste óptimo de los procesos cognitivos y metacognitivos responsables de la estimación subjetiva de éxito. Se afirma, asimismo, que el entrenamiento limitado a tareas verbales específicas también permite disolver el sesgo de subconfianza. Su sugiere que esto ocurre porque el entrenamiento en tareas específicas genera la ocasión de entrenar indirectamente los procesos de la calibración, pero con menor eficacia que el entrenamiento focalizado directamente en la calibración.

Referencias

- Angner, E. (2006). Economists as Experts: Overconfidence in Theory and Practice. *Journal of Economic Methodology*, 13(1), 1-24.
- Barbero García, M.I., Prieto Marañón, P., Suárez Falcón, J.C. y San Luis Costas, C. (2001). Relaciones empíricas entre los estadísticos de la teoría clásica de los tests y los de la teoría de respuesta a los ítems. *Psicothema*, 13(2), 324-329.
- Brenner, L. (2000). Should Observed Overconfidence Be Dismissed as a Statistical Artifact? Critique of Erev, Wallsten, and Budescu (1994). *Psychological Review*, 107(4), 943-946.
- Brenner, L. (2003). A Random Support Model of the Calibration of Subjective Probabilities. *Organizational Behavior and Human Decision Processes*, 90, 87-110.
- Camerer, C.F. y Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review*, 89(1), 306-318.
- Christensen-Szalanski, J.J. y Bushyhead, J.B. (1981). Physicians' Use of Probabilistic Information in a Real Clinical Setting. *Journal of Experimental Psychology: Human perception and performance*, 7, 928-935.
- Clayton, D.E. (2005). Performance Overconfidence: Metacognitive Effects or Misplaced Student Expectations? *Journal of Marketing Education*, 27(2), 122-129.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences. Second Edition*. Hillsdale, NJ: LEA.
- Cortada de Kohan, N. (1999). *Teorías Psicométricas y Construcción de Tests*. Buenos Aires: Lugar Editorial.
- Cortada de Kohan, N. (2003). *BAIRES. Test de Aptitud Verbal*. Madrid: TEA.
- Cortada de Kohan, N. y Macbeth, G. (2006). Los Sesgos Cognitivos en la Toma de Decisiones. *Revista de Psicología UCA*, 3(2), 55-71.
- Dhali, M.K., Hertwig, R. y Hoffrage, U. (2004). The Role of Representative Design in an Ecological Approach to Cognition. *Psychological Bulletin*, 130(6), 959-988.
- Erev, I., Wallsten, T.S. y Budescu, D.V. (1994). Simultaneous Over- and Underconfidence: The Role of Error in Judgment Processes. *Psychological Review*, 101(3), 519-527.
- Finn, B. y Metcalfe, J. (2007). The Role of Memory for Past Test in the Underconfidence With Practice Effect. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 33(1), 238-244.
- Fischhoff, B. (1982). Debiasing. En D. Kahneman, P. Slovic y A. Tversky (Eds.). *Judgment Under Uncertainty: Heuristics and Biases* (pp. 422-444). Cambridge: Cambridge University Press.
- Garriga Trillo, A.J., Villarino, A., González Labra, M.J. y Arnau, M.A. (1994). La Calibración de Juicios Psicofísicos: Estimación de Magnitudes. *Psicothema*, 6(3), 525-532.
- Gigerenzer, G., Hoffrage, U. y Kleinbölting, H. (1991). Probabilistic Mental Models: A Brunswikian Theory of Confidence. *Psychological Review*, 98(4), 506-528.
- Glass, G.V., McGaw, B. y Smith, M.L. (1981). *Meta-Analysis in Social Research*. Thousand Oaks, CA: Sage.
- Grissom, R.J. y Kim, J.J. (2005). *Effect Sizes for Research. A Broad Practical Approach*. Mahwah, NJ: LEA.
- Hambleton, R.K. (1994). Item Response Theory: A Broad Psychometric Framework for Measurement Advances. *Psicothema*, 6(3), 535-556.
- Henrion, M. y Fischhoff, B. (1986). Assessing Uncertainty in Physical Constants. *American Journal of Physics*, 54, 791-797.
- Juslin, P., Winman, A. y Olsson, H. (2000). Naive Empiricism and Dogmatism in Confidence Research: A Critical Examination of the Hard-Easy Effect. *Psychological Review*, 107, 384-396.
- Kahneman, D. (2003). A Perspective on Judgment and Choice. Mapping Bounded Rationality. *American Psychologist*, 58(9), 697-720.
- Kahneman, D. y Tversky, A. (1982). Intuitive Prediction: Biases and Corrective Procedures. En D. Kahneman, P. Slovic y A. Tversky (Eds.). *Judgment Under Uncertainty: Heuristics and Biases* (pp. 414-421). Cambridge: Cambridge University Press.
- Kahneman, D. y Tversky, A. (1996). On the Reality of Cognitive Illusions. *Psychological Review*, 103(3), 582-591.
- Klayman, J., Soll, J.B., Gonzalez-Vallejo, C. y Barlas, S. (1999). Overconfidence: It Depends on How, What and Whom You Ask. *Organizational Behavior and Human Decision Process*, 79, 216-247.
- Koheler, D.J., Brenner, L. y Griffin, D. (2002). The Calibration of Expert Judgment: Heuristics and Biases Beyond the Laboratory. En T. Gilovich, D. Griffin y D. Kahneman (Eds.). *Heuristics and Biases. The Psychology of Intuitive Judgment* (pp. 686-715). Cambridge: Cambridge University Press.
- Lieberman, V. (2004). Local and Global Judgments of Confidence. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 30(3), 729-732.
- Lieberman, V. y Tversky, A. (1993). On the Evaluation of Probability Judgments: Calibration, Resolution, and Monotonicity. *Psychological Bulletin*, 114(1), 162-173.

- Lichtenstein, S., Fischhoff, B. y Phillips, L.D. (1982). Calibration of Probabilities: The State of the Art to 1980. En D. Kahneman, P. Slovic y A. Tversky (Eds.). *Judgment Under Uncertainty: Heuristics and Biases* (pp. 306-334). Cambridge: Cambridge University Press.
- Macbeth, G. (2005). Confidence Calibration and Performance Differences in General Knowledge Tasks. *The Brunswik Society Newsletter*, 20, 11.
- Macbeth, G. (2006). The Effect of Calibration Training on the Underconfidence Bias. *The Brunswik Society Newsletter*, 21, 12.
- Macbeth, G. y Bogiaizian, D. (2007). La estimación subjetiva de éxito en los trastornos de ansiedad. *Revista Argentina de Clínica Psicológica*, 16(2), 143-150.
- Macbeth, G. y Cortada de Kohan, N. (en prensa). El Efecto de Sobreconfianza en Economistas Expertos. *Psico-Logos*.
- Macbeth, G., Cortada de Kohan, N. y Kohan Cortada, A. (2007). El efecto de la experticia matemática en el sesgo de sobreconfianza. *Perspectivas en Psicología*, 4(1), 13-18.
- Macbeth, G., Cortada de Kohan, N., Razumiejczyk, E. y López Alonso, A.O. (2006). Los Sesgos de Sobreconfianza y Subconfianza en Tareas de Conocimientos Generales. *Acta Psiquiátrica y Psicológica de América Latina*, 52(4), 221-226.
- Macbeth, G., Razumiejczyk, E. y Cortada de Kohan, N. (2006). El Sesgo de Sobreconfianza en Tareas Verbales y Matemáticas. *Investigaciones en Psicología UBA*, 11(3), 47-58.
- McGraw, K. y Wong, S. (1992). A common language effect size statistic. *Psychological Bulletin*, 111, 361-365.
- Merkle, E. y Van Zandt, T. (2006). An Application of the Poisson Race Model to Confidence Calibration. *Journal of Experimental Psychology: General*, 135(3), 391-408.
- Murphy, A.H. y Winkler, R.L. (1974). Subjective Probability Forecasting Experiments in Meteorology: Some Preliminary Results. *Bulletin of the American Meteorological Society*, 55, 1206-1216.
- Murphy, A.H. y Winkler, R.L. (1977). Can Weather Forecasters Formulate Reliable Probability Forecasts of Precipitation and Temperature? *National Weather Digest*, 2, 2-9.
- Oskamp, S. (1965). Overconfidence in case-study judgments. *The Journal of Consulting Psychology*, 29, 261-265.
- Simmons, J.P. y Nelson, L.D. (2006). Intuitive Confidence: Choosing Between Intuitive and Nonintuitive Alternatives. *Journal of Experimental Psychology: General*, 135(3), 409-428.
- Svenson, O. (1981). Are We All Less Risky and More Skillful than Our Fellow Drivers? *Acta Psychologica*, 47, 143-148.
- Valera-Espín, A. y Sánchez-Meca, J. (1997). Pruebas de significación y magnitud del efecto: Reflexiones y propuestas. *Anales de psicología*, 13(1), 85-90.
- Wallsten, T.S., Erev, I. y Budescu, D.V. (2000). The Importance of Theory: Response to Brenner (2000). *Psychological Review*, 107(4), 947-949.
- Wilson, T.D., Centerbar, D.B. y Brekke, N. (2002). Mental Contamination and the Debiasing Problem. En T. Gilovich, D. Griffin y D. Kahneman (Eds.). *Heuristics and Biases. The Psychology of Intuitive Judgment* (pp. 185-200). Cambridge: Cambridge University Press.

(Artículo recibido: 6-6-2007; aceptado: 10-4-2008)