

Convocatoria de Ayudas a Proyectos de Innovación y Mejora
Docente 2017-2018

Proyecto:

Detección de mentiras:

Una experiencia práctica en el aula

(ID2017/085)

Memoria Final

COORDINADOR DEL PROYECTO:			
NIF	Nombre y apellidos	E-mail	Teléfono
77833377C	Jaume Masip Pallejà	jmasip@usal.es	Ext. 5645

MIEMBROS DEL EQUIPO DE TRABAJO (sin incluir al coordinador):		
NIF	Nombre y apellidos	E-mail
07850417B	Carmen Herrero Alonso	cherrero@usal.es

Antecedentes

Las actividades constitutivas de este proyecto se llevaron a cabo en la asignatura “Psicología Jurídica” del Grado de Psicología (código de la asignatura: 105459), de la que el coordinador de este proyecto es profesor responsable. Los contenidos abordados mediante tales actividades forman parte de los incluidos en el programa de dicha asignatura. Los dos docentes del equipo son expertos en Psicología Jurídica, siendo de hecho la directora y uno de los miembros del GIR de la USAL “Psicología Jurídica”. También cuentan con una trayectoria investigadora conjunta en torno a la detección de mentiras (temática fundamental en este proyecto) que abarca dos décadas, incluyendo numerosos proyectos financiados y publicaciones internacionales en revistas de impacto. Comparten o han compartido docencia en varias asignaturas, incluyendo algunas del ámbito psicojurídico, y han colaborado con éxito en convocatorias anteriores de Proyectos de Innovación y Mejora Docente de la USAL.

Motivación

Desde que comenzara la adaptación de las distintas titulaciones a los actuales grados, los miembros del equipo docente e investigador que firman la presente memoria se cuestionaron si el alumnado (y los docentes) estaban en las mejores condiciones para afrontar el proceso de enseñanza/aprendizaje. Sobre todo, se plantearon qué se puede hacer para facilitar al estudiante la adopción de ese papel activo que se le exige en los nuevos grados. Tal preocupación estuvo en el origen de la búsqueda de estrategias docentes eficaces y en el desarrollo de recursos y materiales docentes que contribuyeran del mejor modo posible a la adquisición de las competencias establecidas en las titulaciones en las que desarrollan su labor docente.

La consecución de los objetivos en los proyectos de innovación anteriores en los que el equipo ha participado en los últimos años ha permitido ir perfilando y poniendo a prueba distintas actividades participativas en asignaturas de varias titulaciones. Estas actividades se conciben como un recurso necesario para generar en el alumnado el conflicto cognitivo necesario para un aprendizaje significativo. Algunos de los resultados de los proyectos de innovación anteriores se han presentado en congresos nacionales e internacionales, habiendo incluso dado lugar a alguna publicación (JCR).

El presente proyecto supone un avance más en esa dirección, con la introducción de nuevos elementos y el alcance a contenidos y resultados de aprendizaje para los que hasta ahora no habíamos puesto a prueba esta manera de intervenir. Como se puede apreciar en este informe, son los propios alumnos los que, a través de la realización de un “experimento controlado” en el aula (y de todo el trabajo que ello supone para docentes y estudiantes), han generado los datos relativos a los conceptos y procesos objeto de aprendizaje. Sin embargo, en este caso, la actividad que pretendíamos validar no sólo se orientaba a demostrar de modo vivencial un fenómeno o proceso psicológico

de un modo tal que el alumno fuera el sujeto activo, sino también a que dicho alumno fuera capaz de adquirir e interiorizar conceptos complejos de carácter más abstracto.

La actividad ha permitido, además, enfrentar al alumno a posibles falsas creencias y mitos—es decir, conceptos erróneos sobre los contenidos que se explican—. Someter al alumno a una experiencia en el aula que él pueda contrastar con sus creencias erróneas para que advierta la naturaleza falaz de éstas se ha mostrado eficaz en otras ocasiones (considérense los resultados de proyectos de innovación previos de este equipo; por ejemplo, ID/0022, ID9/155 e ID2015/0141). Las concepciones erróneas referentes a la detección de mentiras adquieren una relevancia particular en las asignaturas con contenidos psicojurídicos, y es imperativo corregirlas, ya que pueden estar en la base de una mala praxis en ámbitos aplicados a los que puede dedicarse el alumno en su futuro profesional.

Las experiencias anteriores han revelado que las meras demostraciones en el aula pueden ser insuficientes para contrarrestar las concepciones y adquirir las correctas a no ser que se aborden de una manera específica, controlando y verificando los conceptos que subyacen a los contenidos y fenómenos que se tratan de demostrar. Esto es, no se producirá el cambio conceptual necesario si no se utilizan estrategias específicas que lo puedan suscitar.

Con todo lo expuesto en mente, diseñamos e implementamos la presente actividad con el fin de mejorar los resultados de aprendizaje de los alumnos de la asignatura Psicología Jurídica (105459) del Grado de Psicología.

Objetivos

El **objetivo principal** de este proyecto era que el alumnado adquiriera ciertos conocimientos complejos incluidos en el programa de la asignatura Psicología Jurídica (105459) del Grado de Psicología mediante la participación en una tarea activa que le colocara en el centro de su propio proceso de aprendizaje.

Asimismo, como objetivo más **específico** cabe señalar el que los alumnos generaran, intercambiaran y utilizaran sus propios datos para la adquisición de conocimientos sobre distintos procesos y contenidos fundamentales (en este caso algunos conceptos relacionados con la detección de mentiras). En esencia, se trataba de que experimentaran en su propia persona los procesos y contenidos que son materia de estudio, implicándose en los mismos pero de un modo riguroso y científico. Tal y como la propia ciencia psicológica nos enseña, la información vivida tiene mucho más peso que cualquier otro tipo de información. Todo ello supone que el alumnado, además, incremente su interés, motivación y grado de aprendizaje.

Trasfondo de la Actividad

En detección de mentiras, se diferencia entre el **emisor** (persona que emite una comunicación, que puede ser sincera o mendaz) y el **receptor** (persona que recibe dicha comunicación). La investigación científica ha examinado la *precisión* de los receptores (grado en que estos son capaces de diferenciar entre verdades y mentiras) y su *sesgo de respuesta* (grado en que los receptores tienden a hacer más juicios de verdad que de mentira o viceversa).

La precisión y el sesgo de respuesta pueden depender de cuatro factores. Por parte del **emisor**, pueden depender de la transparencia de éste (que influye sobre la precisión), así como de su credibilidad (que influye sobre el sesgo de respuesta). La *transparencia* es la medida en que el emisor parece sincero al decir la verdad y mentiroso al mentir. Si un emisor es muy transparente, los juicios del receptor serán muy precisos; si es poco transparente, los juicios del receptor serán poco precisos. La *credibilidad* es el grado en el que el emisor da la impresión de sinceridad de manera estable, independientemente de si está diciendo la verdad o mintiendo. Un emisor muy creíble incrementará la tendencia del receptor a hacer juicios de verdad y otro poco creíble aumentará la tendencia del receptor a hacer juicios de mentira.

Por parte del **receptor**, la precisión y el sesgo de respuesta pueden depender de su capacidad (que influye sobre la precisión) y de su credulidad (que influye sobre el sesgo de respuesta). La *capacidad* del receptor es el grado en que éste es capaz de saber si el emisor dice la verdad o miente. Si un receptor es muy capaz, sus juicios serán muy precisos; si es poco capaz, sus juicios serán poco precisos. La *credulidad* es la medida en la que el receptor es confiado y presta credibilidad a los mensajes del emisor independientemente de si son verdaderos o falsos. Un receptor muy crédulo hará muchos juicios de verdad, mientras que un receptor poco crédulo hará muchos juicios de mentira.

Los hallazgos de décadas de investigación muestran lo siguiente:

- (a) Las personas apenas aciertan por encima del nivel de azar al juzgar si otros mienten o dicen la verdad (por ej., Bond y DePaulo, 2006).
- (b) Las personas tienden a pensar que los demás dicen la verdad (sesgo de veracidad), es decir, suelen ser muy crédulas (por ej., Bond y DePaulo, 2006; Levine, Park y McCornack, 1999). En los policías u otros profesionales para quienes detectar mentiras es importante, esta tendencia de respuesta se reduce o incluso se invierte (sesgo de mendacidad; Masip, 2014; Meissner y Kassin, 2002).
- (c) Las personas varían mucho en cuanto a credibilidad: existen personas muy creíbles y personas muy poco creíbles (Bond y DePaulo, 2008; Levine, 2016).
- (d) Las personas varían bastante en transparencia: hay personas bastante transparentes y personas bastante poco transparentes (Bond y DePaulo, 2008).

(e) Las personas varían relativamente poco en credulidad, aunque sí hay cierta variabilidad (Bond y DePaulo, 2008; Levine, 2016).

(f) Las personas apenas varían en capacidad: no existen en la población personas muy capaces de diferenciar entre verdades y mentiras ni personas poco capaces, casi todo el mundo acierta en torno al nivel de azar (Bond y DePaulo, 2008; Levine, 2016).

(g) De manera más general, hay más variabilidad entre emisores (credibilidad y transparencia) que entre receptores (credulidad y capacidad; Bond y DePaulo, 2008; Levine, 2016).

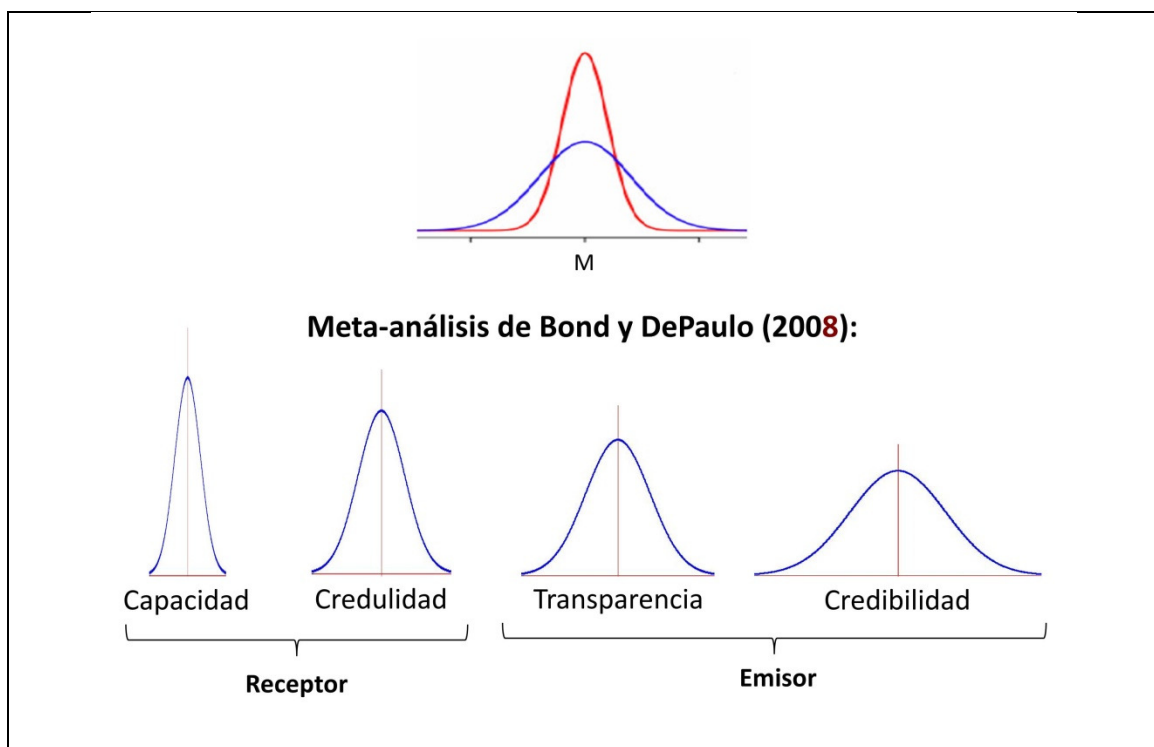


Figura 1. Ilustración gráfica didáctica de los puntos *g* a *f* enumerados en el texto

El objeto de la actividad que se propone era que los alumnos interiorizaran estas nociones mediante un ejercicio participativo en el cual todos ellos deberían actuar como emisores (mintiendo y diciendo la verdad ante sus compañeros de clase) y como receptores (juzgando si cada uno de sus compañeros de clase mentía o decía la verdad). Para cada alumno, se midió su *transparencia* (porcentaje de juicios correctos emitidos por sus compañeros al juzgar la veracidad del alumno en cuestión), su *credibilidad* (porcentaje de juicios de verdad emitidos por sus compañeros), su *capacidad* (porcentaje de juicios correctos emitidos por el alumno en cuestión al juzgar a sus compañeros) y su *credulidad* (porcentaje de juicios de verdad emitidos por ese alumno al juzgar a sus compañeros). En una sesión posterior, se explicó en clase el sentido del ejercicio y su

trasfondo teórico. También se proporcionaron a los alumnos los *resultados globales* obtenidos por el grupo de clase, para ilustrar los puntos *a – g* antes enumerados, así como sus *puntuaciones individuales* para que cada cual pudiera conocer su propio grado de transparencia, credibilidad, capacidad y credulidad.

Para evaluar la efectividad pedagógica del ejercicio, se administró a los alumnos un cuestionario al principio y al final del cuatrimestre. El cuestionario inquiría sobre los aspectos recogidos en los puntos *a – g* anteriores. Esperábamos que las respuestas al cuestionario fueran más precisas después de haber hecho el ejercicio (es decir, al final del cuatrimestre) que antes de haberlo hecho (principio del cuatrimestre).

Metodología y Actividades

Fase 1: Medida del Conocimiento Previo

Con el fin de evaluar la efectividad pedagógica de la actividad, diseñamos un breve cuestionario para medir los conocimientos de los alumnos sobre los contenidos objeto de aprendizaje. Administramos dicho cuestionario tanto antes de llevar a cabo dicha actividad como después de ella. El cuestionario contenía varias secciones:

1. Apartado en el cual se recogía la edad y el sexo de los alumnos, así como la fecha de nacimiento de sus padres (estas fechas se emplearían para poder emparejar los dos cuestionarios que cada participante habría cumplimentado en momentos distintos, manteniendo no obstante su anonimato).
2. Una breve descripción de las nociones de “emisor” y “receptor”, así como de “transparencia”, “credibilidad”, “capacidad” y “credulidad”, similares a las descripciones de estos términos proporcionados unos párrafos más arriba.
3. Un bloque de ocho preguntas en las cuales los alumnos participantes debían señalar en escalas de 1 a 5 la medida en que existían emisores muy transparentes, muy poco transparentes, muy creíbles y muy poco creíbles, y receptores muy capaces, muy poco capaces, muy crédulos y muy poco crédulos. He aquí una pregunta de ejemplo:

¿Crees que existen emisores muy transparentes (es decir, personas a las que casi siempre se les nota si mienten o dicen la verdad)?

No				Sí
1	2	3	4	5

4. Un bloque de cuatro preguntas en las cuales el participante debía señalar cuán transparente (1) o poco transparente (5), creíble (1) o poco creíble (5), capaz (1) o poco capaz (5), y crédulo (1) o poco crédulo (5) se consideraba a sí mismo. La siguiente era una de estas preguntas, que incluimos aquí como ejemplo:

¿Cómo te definirías a ti misma/o?

Soy una persona:

Muy transparente (casi siempre se me nota si miento o digo la verdad)

Muy poco transparente (casi nunca se me nota si miento o digo la verdad)

1

2

3

4

5

Esperábamos que al final del cuatrimestre, tras haber llevado a cabo la tarea, los participantes tendrían nociones más ajustadas (al menos en lo referente a otras personas) que al principio del mismo.

El cuestionario se administró por primera vez al inicio de la primera clase del cuatrimestre correspondiente a la asignatura. Lo cumplimentaron, de manera voluntaria, 41 estudiantes (29 mujeres y 12 hombres, $M_{\text{edad}} = 22$ años, $DT = 1,80$, rango: 20-30).

Fase 2: Ejercicio Participativo

El ejercicio central de este proyecto se realizó en horario de clase. Tomaron parte, de manera voluntaria y tras la firma de un formulario de consentimiento informado, las 37 mujeres y los 13 hombres ($M_{\text{edad}} = 22$ años, $DT = 1,66$, rango: 20-30 años, todos estudiantes de 4º curso del Grado de Psicología) que ese día habían acudido a clase.

Se entregó a cada participante una etiqueta con un número identificativo que tendrían que escribir en el encabezado de cada cuestionario que cumplimentaran. Además, se les indicó que debían guardar ese número a fin de poder consultar los resultados individuales que obtendrían en la tarea cuando estos estuvieran disponibles.

El primer cuestionario contenía las siguientes ocho preguntas. Los participantes debían responderlas con sinceridad por escrito, aportando además un breve razonamiento de su respuesta:

1. ¿Cuál es el mejor lugar al que has ido de vacaciones y por qué te gustó tanto?
2. ¿Cuál ha sido la peor clase a la que has asistido y por qué fue tan mala?
3. ¿En qué trabajan tus padres?
4. ¿Cuál te gusta más, primavera, verano, otoño o invierno y por qué?
5. ¿Qué es lo que te da más miedo y por qué?
6. ¿Qué vacaciones te gustan más y por qué? Navidad, verano o Semana Santa.
7. ¿Cuál es tu película favorita y por qué?
8. ¿Cuál es tu grupo o cantante de música favorito?

Mediante la administración de este cuestionario pudimos conocer la verdad acerca de los temas sobre los que, en breve, deberían mentir los estudiantes. Esto nos permitiría constatar que, efectivamente, seguían nuestras instrucciones de mentir o decir la verdad.

Después de recoger los cuestionarios cumplimentados, llevamos a cabo la tarea central del ejercicio. En concreto, se llamaba a cada alumno (a partir del número de su etiqueta) para que se personara al frente de la clase. Se le colocaba un micrófono para poder grabar sus respuestas y se le entregaba un cuestionario. Este era idéntico al que acababa de responder, pero cuatro preguntas estaban señaladas para que mintiera cuando se le formularan en alto (estas preguntas estaban en negrita e iban seguidas de la palabra “MENTIRA” escrita en mayúscula y negrita). Las preguntas señaladas eran distintas para cada participante y se habían determinado previamente mediante un software de extracción de números aleatorios.¹

Los receptores (es decir, todos los alumnos participantes) recibieron una hoja de respuesta en la cual, para cada participante (nº 1, nº 2, etc.), debían anotar si les parecía que la respuesta a cada pregunta era la verdad (“V”) o una mentira (“M”). No se informó a los participantes del número de verdades y de mentiras que iba a contar cada emisor (en realidad fueron siempre cuatro verdades y cuatro mentiras).

Un alumno colaborador fue formulando cada pregunta en voz alta a cada alumno, y éste debía ir respondiendo, mintiendo en las preguntas señaladas y diciendo la verdad en las restantes. Después de cada respuesta, los asistentes señalaban en el cuestionario si la respuesta les parecía verdadera o falsa. Al mismo tiempo, una asistente (sentada en la mesa del profesor) iba comparando las respuestas de cada emisor con las ofrecidas previamente en el cuestionario inicial, registrando en cuáles mentía. El objeto de esta última tarea era constatar que, efectivamente, los participantes mentían ante las preguntas designadas y no ante otras. La grabación de las respuestas permitiría revisar este punto con mayor minuciosidad posteriormente.

Se indicó a los alumnos que en caso de conocer alguna respuesta debido una relación de amistad con el emisor, escribieran un asterisco junto a dicha respuesta. Las respuestas seguidas de un asterisco se desestimaron con el objeto de no alterar los resultados, pues en este caso la elevada precisión no se debería a una gran transparencia del emisor ni a una gran capacidad del detector, sino a una “variable extraña”: el conocimiento íntimo entre emisor y receptor.

Nótese que todos los alumnos ($n = 50$) actuaron como emisores y como receptores (juzgando la veracidad de sus 49 compañeros). Esto nos permitiría calcular la *transparencia* de cada alumno (porcentaje de juicios correctos efectuados por sus

¹ Sobre un máximo previsto de 57 participantes, extrajimos 57 conjuntos de cuatro números aleatorios entre 1 y 8 y sin repetición. Las cuatro preguntas a las que debía mentir el Emisor 1 serían las preguntas cuyos números correspondieran con los cuatro números extraídos en el primer conjunto, las cuatro preguntas a las que debía mentir el Emisor 2 serían aquellas cuyos números se correspondieran con los cuatro números extraídos en el segundo conjunto y así sucesivamente.

compañeros al juzgar si sus respuestas eran verdaderas o falsas), su *credibilidad* (porcentaje de juicios de verdad efectuados por sus compañeros al juzgar si sus respuestas eran verdaderas o falsas), su *capacidad* (el porcentaje de juicios correctos efectuados por él al juzgar la veracidad de las respuestas de sus compañeros) y su *credulidad* (el porcentaje de juicios de verdad efectuados por él al juzgar la veracidad de las respuestas de sus compañeros).

Una vez finalizada la tarea, se recogieron los cuestionarios y se informó a los alumnos de que en un momento posterior del curso se dedicaría una clase a exponer los resultados obtenidos. Los datos se analizaron con Excel y SPSS.

Fase 3: Análisis y Feedback

Recogimos un total de 19.600 juicios de veracidad (50 receptores x 49 emisores juzgados por cada receptor x 8 respuestas por emisor). A la hora de analizar los resultados, desestimamos los 229 juicios (1,17% del total) seguidos de un asterisco. Además, detectamos que 11 de los 50 emisores no siguieron estrictamente las instrucciones, confundiendo y mintiendo o diciendo la verdad a destiempo. Esto hizo que la proporción de verdades y mentiras no fuera del 50%, sino del 49% para las verdades y del 51% para las mentiras.

Esperábamos replicar los resultados meta-analíticos de Bond y DePaulo (2006, 2008) y los del estudio reciente de Levine (2016), que en esencia muestran los puntos siguientes (que ya especificamos unas páginas más atrás): (a) las personas apenas aciertan por encima del nivel del nivel de azar al juzgar si otros mienten o dicen la verdad; (b) la gente suele ser muy crédula, pensando que los demás dicen la verdad (sesgo de veracidad); (c) las personas varían más en credibilidad y en transparencia (dimensiones del emisor) que en credulidad y en capacidad (dimensiones del receptor). En concreto, la variabilidad en capacidad se ha mostrado ínfima: apenas hay personas muy capaces o muy poco capaces de separar entre verdades y mentiras, casi todo el mundo acierta en torno al nivel de azar.

Nuestros resultados están en línea con el punto *a*: el porcentaje de juicios correctos fue del 55% (IC del 95% \pm 0,95%), superior al azar (50%), $t(49) = 9,97$, $p < 0,001$, y casi idéntico al promedio del 54% del meta-análisis de Bond y DePaulo (2006), $t(49) = 1,49$, $p = .142$. También están en consonancia con el punto *b*: el porcentaje de juicios de verdad fue del 63% (IC del 95% \pm 2,52%), muy superior al 50% esperado por azar, $t(49) = 10,69$, $p < 0,001$, e incluso superior al porcentaje meta-analítico del 55% encontrado por Bond y DePaulo (2006), $t(49) = 6,69$, $p < 0,001$ (sin embargo, el 23% de las 207 comparaciones relevantes del meta-análisis de Bond y DePaulo, 2006, tenían un sesgo de veracidad igual o más extremo que el hallado por nosotros).

En cuanto al punto *c*, la Figura 2 muestra la distribución de las puntuaciones en cada una de las cuatro dimensiones relevantes. Podemos ver con claridad que, tal como

esperábamos, la variabilidad fue mayor en las dimensiones del emisor (véanse cuadrantes inferiores de la Figura 2; para la credibilidad: $S^2 = 68\%$; para la transparencia: $S^2 = 70\%$) que en las del receptor (cuadrantes superiores de la Figura 2), con una excepción: una gran variabilidad en credulidad, $S^2 = 78\%$. En concreto, el

<p>Credulidad (juicios de verdad) $M = 63\%$ $S^2 = 78\%$</p> <p>1 1 2 2 3 3 4 4 4 9 5 0001334 5 6788999 6 000012222344 6 555558888999 7 333331 7 89 8 34 8 9 9</p>	<p>Capacidad (aciertos) $M = 55\%$ $S^2 = 11\%$</p> <p>1 1 2 2 3 3 4 4 689 5 011223333333333344444444444444 5 56666777778 6 0001113 6 7 7 8 8 9 9</p>
<p>Credibilidad $M = 63\%$ $S^2 = 68\%$</p> <p>1 1 2 2 3 3 4 4 7889 5 0134 5 55567889 6 002444 6 5566666667899 7 0111111224 7 5567 8 8 9 9</p>	<p>Transparencia $M = 55\%$ $S^2 = 70\%$</p> <p>1 1 2 2 3 4 3 4 00244 4 66888999 5 11112223444 5 66677777888 6 00112331 6 5688 7 14 7 8 8 9 9</p>

Figura 2. Resultados de la tarea

porcentaje de juicios de verdad efectuados por los receptores abarcó entre el 44% y el 84%, y además la distribución fue bastante plana (primer cuadrante de la Figura 2). En cambio, de acuerdo con lo esperado, la variabilidad en capacidad fue nimia, $S^2 = 11\%$. En la Figura 2 (segundo cuadrante) podemos apreciar que el rango fue del 46% al 63%,

y que de 50 participantes, hubo 29 que acertaron entre el 50 y el 54% de las veces y 40 (un 80%) que acertaron entre el 50 y el 58% de las veces. Claramente, en nuestra muestra no hubo ni buenos ni malos detectores.

Los resultados se expusieron a los alumnos durante una sesión teórica doble. La presentación incluyó la diferenciación entre emisor y receptor, así como la explicación de las cuatro dimensiones de interés (credibilidad, transparencia, credulidad y capacidad), los resultados más relevantes de los meta-análisis previos de Bond y DePaulo (2006, 2008) y del estudio de Levine (2016), ilustrados con gráficos muy claros, y por último la presentación detallada de los resultados obtenidos en nuestro ejercicio, que comparamos con los de Bond y DePaulo (2006, 2008) y de Levine (2016). Después de haber impartido esta sesión, se subió a Studium un documento en el cual figuraba el número de cada participante, seguido de sus puntuaciones en capacidad, credulidad, transparencia y credibilidad (Figura 3).

Capacidad: Juicios que yo acerté (porcentaje)
 Credulidad: Juicios de verdad que yo hice (porcentaje)
 Transparencia: Porcentaje de mis comunicaciones que los demás juzgaron correctamente
 Credibilidad: Porcentaje de mis comunicaciones que los demás consideraron que eran verdad

Participante Número	Capacidad	Credulidad	Transparencia	Credibilidad
1	48	60	58	57
2	54	59	66	72
6	53	64	52	48
8	56	50	48	74
9	54	53	58	70
11	50	58	49	54
13	51	60	53	71
15	63	65	54	66
17	57	69	57	51
18	54	50	46	66

Figura 3. Inicio del informe de *feedback* personalizado

Fase 4: Medida del Conocimiento Final

Con el objeto de medir la efectividad pedagógica de este ejercicio docente, el día del examen, antes de proceder al mismo, se administró a los alumnos el mismo cuestionario que el primer día del cuatrimestre, aunque se añadieron dos preguntas: “¿Participaste en el ejercicio sobre detección de mentira que hicimos en clase? Si/No” y “¿Estuviste en clase el día en que se explicó el ejercicio? Si/No”.

El número de estudiantes que cumplimentó el cuestionario en ambas ocasiones fue 35. De ellos, 24 manifestaron que, además, habían participado en el ejercicio de clase y habían asistido a la sesión de *feedback*. Los cálculos del apartado siguiente se

efectuaron sobre estos 24 participantes (18 mujeres y seis hombres; edad al final del cuatrimestre: $M = 22$ años, $DT = 2,33$, rango: 20-31).

Efectividad Pedagógica de la Tarea

La efectividad pedagógica de las actividades realizadas se midió comparando formalmente las respuestas dadas al cuestionario al principio del cuatrimestre y al final del mismo por los 24 alumnos que, además de cumplimentar el cuestionario en ambas ocasiones, habían asistido a las dos sesiones relevantes en clase (el ejercicio práctico y la sesión posterior de *feedback*). Cabía esperar que las respuestas de estos alumnos fueran más acertadas al final del cuatrimestre que al principio del mismo.

Primer Bloque de Preguntas

Nótese que en el primer bloque de preguntas los participantes tuvieron que indicar en escalas de 1 (*No*) a 5 (*Sí*) la medida en que creían que existían emisores muy transparentes (primera pregunta) y emisores muy poco transparentes (segunda pregunta). Esto significa que había dos preguntas distintas referidas a la variabilidad en cuanto a transparencia. Con el fin de aunar las dos medidas para obtener una puntuación en una única dimensión, se aplicó la siguiente fórmula:

$$(\text{Puntuación dada a la primera pregunta} - 1) + (\text{Puntuación dada a la segunda pregunta} - 1)$$

Esto creó una variable combinada de variabilidad que abarcaba desde 0 (mínima variabilidad en transparencia) a 8 (máxima variabilidad en transparencia). Por ejemplo, un participante que hubiera señalado 1 (mínima puntuación) en ambas preguntas obtendría un 0 en la escala: $(1-1) + (1-1) = 0 + 0 = 0$. Otro que hubiera señalado 5 (máxima puntuación) en ambas escalas obtendría un 8: $(5-1) + (5-1) = 4 + 4 = 8$. Se aplicó la misma fórmula también a las medidas de credibilidad, capacidad y credulidad.

Llevamos a cabo un Análisis de Varianza (ANOVA) de medidas repetidas en el cual introdujimos el Momento de Medición (inicio vs. final del cuatrimestre) y el Constructo Medido (transparencia, credibilidad, capacidad, credulidad) como variables independientes y las puntuaciones en la escala combinada de variabilidad como medida dependiente. Obtuvimos un efecto principal significativo del momento de medición, $F(1, 23) = 7,38$, $p = 0,012$, $\eta^2_{\text{parcial}} = 0,24$, indicativo de mayores valores globales al inicio del cuatrimestre, $M = 5,69$, $DT = 0,83$, que al final del mismo, $M = 4,86$, $DT = 1,44$, y un efecto principal significativo del constructo medido, $F(3, 69) = 11,02$, $p < 0,001$, $\eta^2_{\text{parcial}} = 0,32$, que refleja el hecho de que la puntuación promedio de capacidad ($M = 4,00$, $DT = 1,57$), era estadísticamente menor que las de credulidad ($M = 6,00$, $DT = 1,35$), transparencia ($M = 5,56$, $DT = 1,42$) y credibilidad ($M = 5,54$, $DT = 1,42$), que no difirieron significativamente entre sí. Pero lo que interesa de cara a la efectividad pedagógica de la tarea es la interacción de ambas variables, que fue significativa, $F(3, 69) = 8,86$, $p < 0,001$, $\eta^2_{\text{parcial}} = 0,28$. Las medias y las desviaciones típicas se indican en la Tabla 1.

Tabla 1

Puntuaciones Medias (Desviaciones Típicas entre Paréntesis) en las Variables del Emisor y del Receptor al Principio y al Final del Cuatrimestre (Primer Bloque de Preguntas)

	Inicio del Cuatrimestre	Final del Cuatrimestre
<i>Emisor</i>		
Transparencia	5,83 ^a (1,55)	5,29 ^a (2,33)
Credibilidad	5,25 ^a (1,33)	5,83 ^a (2,12)
<i>Promedio</i>	5,54 ^a (1,16)	5,56 ^a (1,84)
<i>Receptor</i>		
Capacidad	5,17 ^a (1,71)	2,83 ^b (2,01)
Credulidad	6,50 ^a (1,02)	5,50 ^b (2,21)
<i>Promedio</i>	5,83 ^a (1,14)	4,17 ^b (1,85)

Nota: En cada fila, las medias con un superíndice distinto difieren significativamente.

Al inicio del cuatrimestre, todas las puntuaciones estaban significativamente por encima del valor medio (4) de la escala ($ps \leq 0,003$). Hicimos comparaciones *post-hoc* que muestran que al inicio del cuatrimestre la media referente a capacidad era menor que la referente a credulidad ($p = 0,004$), pero no que las correspondientes a transparencia ni credibilidad, que no diferían entre sí. La media para credulidad difería de la de capacidad (como ya hemos indicado) y de la de credibilidad (para ambas comparaciones, $p = 0,004$), pero no difería de la media para transparencia.

Puesto que el mensaje principal de la tarea y de la explicación en clase había sido que la variabilidad en características del receptor (sobre todo en capacidad) era escasa y menor que la variabilidad en características del emisor, esperábamos encontrar valores menores en credulidad y sobre todo en capacidad al final del cuatrimestre que al principio del mismo. Por el contrario, no esperábamos encontrar tales reducciones en transparencia y credibilidad. Según se aprecia en la Tabla 1, esto es justamente lo que hallamos. Mientras que las medias para transparencia y credibilidad no variaron significativamente a lo largo del cuatrimestre: $F(1, 23) = 0,93$, $p = 0,345$, $\eta^2_{\text{parcial}} = 0,04$, y $F(1, 23) = 1,85$, $p = 0,188$, $\eta^2_{\text{parcial}} = 0,07$, respectivamente, las medias para capacidad y para credulidad fueron significativamente menores al final del cuatrimestre que al principio. El descenso fue, según lo esperado, mucho más acusado para la capacidad, $F(1, 23) = 31,52$, $p < 0,001$, $\eta^2_{\text{parcial}} = 0,58$, que para la credulidad, $F(1, 23) = 5,31$, $p = 0,031$, $\eta^2_{\text{parcial}} = 0,19$ (Tabla 1). De hecho, mientras que al final del cuatrimestre los valores promedio para la transparencia, la credibilidad y la credulidad seguían estando significativamente por encima del valor central (4) de la escala ($ps \leq 0,012$), la media correspondiente a la capacidad estaba significativamente por debajo de dicho valor central ($p = 0,009$), ubicándose por lo tanto mucho más cerca de 0 (variabilidad mínima) que de 8 (variabilidad máxima).

Nótese, sin embargo, que según indican las pruebas *post-hoc*, incluso al final del cuatrimestre la media para la credulidad no fue significativamente inferior que para la transparencia ni para la credibilidad. Estos efectos nulos pueden quizás achacarse a que en el ejercicio hecho en clase la variabilidad en credulidad fue inesperadamente elevada y comparable a la variabilidad de la credibilidad y de la transparencia. Aun así, hicimos un ANOVA Momento de Medición (inicio vs. final del cuatrimestre) x Rol (variables del emisor vs. variables del receptor). De nuevo emergió el efecto significativo del momento de medición, $F(1, 23) = 7,38, p = 0,012, \eta^2_{\text{parcial}} = 0,24$. El efecto principal del rol no fue significativo, $F(1, 23) = 2,91, p = 0,102, \eta^2_{\text{parcial}} = 0,11$. Sin embargo, tal como esperábamos, la interacción sí fue significativa, $F(1, 23) = 12,05, p = 0,002, \eta^2_{\text{parcial}} = 0,34$. Las medias y las desviaciones típicas se muestran en las filas de la Tabla 1 llamadas “promedio”. El descenso a lo largo del cuatrimestre fue fuerte y significativo para las variables del receptor, $F(1, 23) = 19,83, p < 0,001, \eta^2_{\text{parcial}} = 0,46$, pero fue literalmente nulo para las del emisor, $F(1, 23) = 0,003, p = 0,959, \eta^2_{\text{parcial}} = 0,003$. Al principio del cuatrimestre, no había diferencias entre las puntuaciones medias correspondientes a las variables del receptor, $M = 5,83, DT = 1,14$, y las del emisor, $M = 5,54, DT = 1,16, F(1, 23) = 0,81, p = 0,337, \eta^2_{\text{parcial}} = 0,03$. Sin embargo, de acuerdo con lo esperado, al final del cuatrimestre la media correspondiente al receptor, $M = 4,17, DT = 1,85$, fue significativamente menor que la correspondiente al emisor, $M = 5,56, DT = 1,84, F(1, 23) = 8,76, p = 0,007, \eta^2_{\text{parcial}} = 0,28$ (Tabla 1). Todos estos análisis demuestran que el ejercicio fue efectivo en términos pedagógicos.

Segundo Bloque de Preguntas

Sabido es en Psicología que a las personas nos cuesta menos percibir y preocuparnos de los sesgos ajenos que de los propios (Pronin, 2007). Este fenómeno, etiquetado como “punto ciego al sesgo” (*bias black spot*), es muy insidioso y resistente ante la formación orientada a eliminarlo (Pronin, Lin y Ross, 2002). Esto sugiere que, en lo que a nuestra actividad pedagógica se refiere, quizás sea más fácil para nuestros alumnos aprender que *la gente en general* varía poco en transparencia, credibilidad, capacidad y credulidad, que cambiar *la propia visión de uno mismo* sobre la medida en que uno es transparente, creíble, capaz o crédulo. Es decir, es concebible que aunque el ejercicio realizado haya logrado que los alumnos entiendan que hay menor variabilidad en las características del emisor (sobre todo en capacidad) que en las del receptor, sin embargo ello no haya cambiado la visión que cada alumno tiene de sí mismo como alguien más o menos capaz o crédulo que la gente en general. Es por esta razón que el segundo bloque de preguntas se refería a la visión que cada participante tenía sobre *su propia* transparencia, credibilidad, capacidad y credulidad.

Dado el modo en que estaba construida la escala empleada en estas cuatro preguntas, la no-variabilidad vendría reflejada por el punto medio (3). Tanto el 2 como el 4 en la escala reflejarían el mismo grado de desviación (un punto) respecto a este punto medio, y lo mismo podría decirse del 1 y del 5 (que reflejarían dos puntos de desviación). En

consecuencia, recodificamos el valor 3 como 0, los valores 2 y 4 como 1 y los valores 1 y 5 como 2. Con esto obtuvimos una medida que abarcaba desde 0 (nula variabilidad) hasta 2 (variabilidad máxima). Si la tarea fue efectiva al cambiar la visión de los alumnos sobre sí mismos, esperaríamos que al final del cuatrimestre los valores para las características del emisor (y sobre todo para la capacidad) fueran significativamente menores (y cercanas a 0) en esta variable que al principio del cuatrimestre, al tiempo que deberían ser menores que las características del receptor.

Hicimos dos ANOVAs análogos a los descritos anteriormente, pero en este caso la variable dependiente la constituyeron las puntuaciones en la nueva escala. En el ANOVA Momento de Medición x Constructo Medido, el efecto principal del momento de medición no fue significativo, $F(1, 23) = 0,09, p = 0,769, \eta^2_{\text{parcial}} = 0,004$, pero el efecto del constructo medido sí lo fue,² $F(2,39, 54,86) = 5,22, p = 0,006, \eta^2_{\text{parcial}} = 0,19$. El valor máximo fue el correspondiente a la credibilidad, $M = 1,08, DT = 0,52$, y el mínimo el correspondiente a la capacidad, $M = 0,63, DT = 0,45$. Los valores referidos a transparencia, $M = 0,80, DT = 0,45$, y a credulidad, $M = 0,77, DT = 0,47$, no difirieron ni entre sí, ni respecto a la credibilidad, ni respecto a la capacidad, pero la diferencia entre credibilidad y capacidad sí fue significativa ($p = 0,012$). Pero, de nuevo, el interés recae en la interacción, que fue significativa, $F(3, 69) = 3,13, p = 0,031, \eta^2_{\text{parcial}} = 0,120$. Las medias y sus correspondientes desviaciones típicas se muestran en la Tabla 2.

Tabla 2

Puntuaciones Medias (Desviaciones Típicas entre Paréntesis) en las Variables del Emisor y del Receptor al Principio y al Final del Cuatrimestre (Segundo Bloque de Preguntas)

	Inicio del Cuatrimestre	Final del Cuatrimestre
<i>Emisor</i>		
Transparencia	0,77 ^a (0,51)	0,83 ^a (0,56)
Credibilidad	1,25 ^a (0,61)	0,92 ^a (0,72)
<i>Promedio</i>	<i>1,01^a (0,41)</i>	<i>0,87^a (0,45)</i>
<i>Receptor</i>		
Capacidad	0,50 ^a (0,51)	0,75 ^a (0,68)
Credulidad	0,71 ^a (0,62)	0,83 ^a (0,56)
<i>Promedio</i>	<i>0,60^a (0,42)</i>	<i>0,79^a (0,44)</i>

Nota: Se observa que, en todas las filas, las dos medias tienen el mismo superíndice, lo cual indica que el efecto de la variable momento de medición no fue significativa para ninguna de las medidas dependientes.

Pese a que el efecto de la interacción fue significativo, la diferencia entre los valores obtenidos al final del cuatrimestre y los obtenidos al principio del mismo no fue significativa en ningún caso: para transparencia, $F(1, 23) = 0,26, p = 0,612, \eta^2_{\text{parcial}} =$

² La prueba de esfericidad fue significativa para esta variable, de modo que empleamos la prueba de Greenhouse-Geisser.

0,01; para credibilidad, $F(1, 23) = 4,00, p = 0,057, \eta^2_{\text{parcial}} = 0,15$; para capacidad, $F(1, 23) = 2,38, p = 0,137, \eta^2_{\text{parcial}} = 0,09$; para credulidad, $F(1, 23) = 0,68, p = 0,417, \eta^2_{\text{parcial}} = 0,03$.³ Vemos, pues, que la reducción esperada para capacidad y para credibilidad no se dio.

El ANOVA Momento de Medición x Rol arrojó un efecto significativo del rol, $F(1, 23) = 20,97, p < 0,001, \eta^2_{\text{parcial}} = 0,48$: las puntuaciones fueron mayores para el emisor, $M = 0,94, DT = 0,35$, que para el receptor, $M = 0,70, DT = 0,32$. La interacción también fue significativa, $F(1, 23) = 6,72, p = 0,016, \eta^2_{\text{parcial}} = 0,23$. Las medias y las desviaciones típicas se muestran en las filas de la Tabla 2 llamadas “promedio”. No hubo diferencias significativas entre la puntuación promedio al final del cuatrimestre y al principio del mismo ni para las variables del emisor, $F(1, 23) = 1,88, p = 0,183, \eta^2_{\text{parcial}} = 0,08$, ni para las del receptor, $F(1, 23) = 2,62, p = 0,119, \eta^2_{\text{parcial}} = 0,10$.⁴

Discusión y Conclusiones

La actividad pedagógica implementada en las clases de la asignatura Psicología Jurídica (105459) del grado en Psicología entronca con algunos de los proyectos de innovación docente previos de este equipo (ID/0022, ID9/155 e ID2015/0141) en el sentido de crear un conflicto cognitivo en el alumno al someterlo a una experiencia en el aula que pueda contrastar con sus creencias erróneas para que advierta la naturaleza falaz de éstas. El objeto principal era que el alumnado adquiriera ciertos conocimientos complejos incluidos en el programa de la asignatura mediante la participación en una tarea activa que le colocara en el centro de su propio proceso de aprendizaje. Los resultados del primer bloque de preguntas del cuestionario empleado para examinar la efectividad pedagógica del ejercicio muestran que este objetivo se ha cumplido. Efectivamente, como resultado de dicho ejercicio los alumnos redujeron significativamente las puntuaciones de variabilidad otorgadas a ambas variables del receptor pero no a las variables del emisor. Además, la reducción fue mayor para la capacidad que para la credulidad. Esto demuestra la eficacia de las actividades realizadas.

Sin embargo, si bien los alumnos fueron capaces de adquirir los conocimientos cruciales, no se los aplicaron a sí mismos, según indican los análisis del segundo bloque de preguntas. Estos análisis no muestran la reducción esperada en las variables del receptor. Este hecho resulta llamativo, puesto que se proporcionó a los alumnos *feedback* individualizado al poner a su disposición en *Studium* el listado con las medidas de transparencia, credibilidad, capacidad y credulidad de cada uno de ellos. Sin embargo, no podemos precisar si los alumnos realmente consultaron esta información ni

³ Lo que refleja la interacción es que si bien al final del cuatrimestre ninguna de las cuatro medias difería de las otras, al principio del mismo los valores en credibilidad eran significativamente superiores a los correspondientes a las otras tres variables, que no diferían entre sí.

⁴ Lo que muestra la interacción es que al principio del cuatrimestre el valor promedio para las variables del emisor era significativamente mayor que el correspondiente a las variables del receptor, $F(1, 23) = 29,37, p < 0,001, \eta^2_{\text{parcial}} = 0,56$; sin embargo, esta diferencia no alcanzó la significación estadística al final del cuatrimestre, $F(1, 23) = 0,89, p = 0,357, \eta^2_{\text{parcial}} = 0,04$ (Tabla 2).

si recordaban el número que se les había asignado el día de la actividad práctica. Sin este número no podían conocer su ejecución individualizada. En cualquier caso, la verdad es que el efecto nulo deja de sorprender si consideramos lo difícil que es reducir el *bias blind spot*, como ha mostrado la investigación en Psicología (Pronin *et al.*, 2002).

Aunque quizás, después de todo, nuestros estudiantes no estuvieran tan sesgados. Efectivamente, un fenómeno llamativo y que puede haber contribuido a la ausencia (en el segundo bloque de preguntas) de los efectos esperados es que, aunque hubiera pocas diferencias significativas (posiblemente debido al limitado tamaño de la muestra, que resultó en desviaciones típicas grandes), el patrón de respuesta, tanto al inicio del cuatrimestre como al final, se ajusta considerablemente al detectado empíricamente por Bond y DePaulo (2008): máxima variabilidad para la credibilidad, seguida de la transparencia, seguida de la credulidad y mínima variabilidad para la capacidad (véase la Tabla 2). Este sorprendente hallazgo sugiere que es posible que los alumnos tengan nociones bastante exactas y robustas sobre sus capacidades y sesgos de respuesta a la hora de emitir verdades y mentiras, así como a la hora de juzgar la credibilidad de las verdades y mentiras de otras personas. Si es así, no había razón para que corrigieran sus respuestas en línea con lo enseñado durante el cuatrimestre: sus respuestas ya eran consistentes con lo enseñado al principio del cuatrimestre.

En definitiva, consideramos que la tarea diseñada y utilizada es interesante y divertida (así nos lo manifestaron verbalmente algunos alumnos) y que cumple el objetivo previsto de interiorizar de un modo ameno y participativo ciertos conocimientos complejos de la asignatura. Tenemos la intención de mandar un breve artículo describiendo dicha tarea y su efectividad a la revista *Teaching of Psychology*, indexada en los *Journal Citation Reports*.

Referencias

- Bond, C. F., Jr. y DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, 10, 214-234. doi:10.1207/s15327957pspr1003_2
- Bond, C. F., Jr. y DePaulo, B. M. (2008). Individual differences in judging deception: Accuracy and bias. *Psychological Bulletin*, 134, 477-492. doi:10.1037/0033-2909.134.4.477
- Levine, T. R. (2016). Examining sender and judge variability in honesty assessments and deception detection accuracy: Evidence for a transparent liar but no evidence of deception general ability. *Communication Research Reports*, 33, 188-194. doi: 10.1080/08824096.2016.1186629
- Levine, T. R., Park, H. S. y McCornack, S. A. (1999). Accuracy in detecting truths and lies: Documenting the "veracity effect". *Communication Monographs*, 66, 125-144. doi: 10.1080/03637759909376468
- Masip, J. (2014). Investigator bias. En T. Levine (Ed.), *Encyclopedia of deception* (Vol. 2, pp. 538-540). Thousand Oaks, CA: Sage.
- Meissner, C. A. y Kassin, S. M. (2002). "He's guilty!": Investigator bias in judgments of truth and deception. *Law and Human Behavior*, 26, 469-480. doi:10.1023/A:1020278620751
- Pronin, E. (2007). Perception and misperception of bias in human judgment. *Trends in Cognitive Sciences*, 11, 37-43. doi:10.1016/j.tics.2006.11.001
- Pronin, E., Lin, D. Y. y Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin*, 28, 369-381. doi:10.1177/0146167202286008