



TESIS DOCTORAL

AÑO: 2015

TÍTULO:

PREDICCIONES DE RECONOCIMIENTO EPISÓDICO

AUTOR:

CRISTÓBAL ARROYO CABRERA

**LICENCIADO EN FILOSOFÍA Y CIENCIAS
DE LA EDUCACIÓN. SECCIÓN: PSICOLOGÍA**

CENTRO ACADÉMICO:

DEPARTAMENTO DE PSICOLOGÍA BÁSICA I

FACULTAD DE PSICOLOGÍA. UNED

DIRECTOR:

MARCOS RUIZ RODRÍGUEZ



DEPARTAMENTO DE PSICOLOGÍA BÁSICA I

**TESIS DOCTORAL:
*PREDICCIONES DE RECONOCIMIENTO EPISÓDICO***

CRISTÓBAL ARROYO CABRERA

**LICENCIADO EN FILOSOFÍA Y CIENCIAS
DE LA EDUCACIÓN. SECCIÓN: PSICOLOGÍA**

**DIRECTOR:
MARCOS RUIZ RODRÍGUEZ**

AGRADECIMIENTOS

Me gustaría expresar mi más sincero agradecimiento a tres personas que me han ayudado en la realización de este trabajo.

En primer lugar, quiero dar las gracias a mi director de tesis Marcos Ruiz por todo lo que me ha enseñado, además de su dedicación y paciencia en los últimos años.

En segundo lugar, quiero agradecer a Antonio Gómez Íñiguez sus ánimos y apoyo moral para que siguiera adelante.

Y por último, quiero agradecer a Rosa María Banda Jiménez su inestimable ayuda en la corrección y revisión de los datos.

ÍNDICE

1.- Predicciones de reconocimiento episódico.....	5
2.- Aproximaciones teóricas en metamemoria.....	7
2.1.- Disociación entre las predicciones y el rendimiento memorístico.....	9
2.2.- Fuentes de información en metamemoria.....	10
3.- Procesamiento Apropriado para la Transferencia.....	14
4.- Supervisión Apropriada para la Transferencia (SAT).....	15
5.- Críticas a la hipótesis SAT.....	17
6.- Planteamiento del problema y formulación de hipótesis.....	19
7.- Objetivos de investigación.....	26
8.- Medida de la precisión de los Juicios de Aprendizaje.....	28
9.- Sección Experimental.....	30
9.1.- Experimento 1.....	31
9.2.- Experimento 2.....	44
9.3.- Experimento 3.....	54
10.- Discusión general y conclusiones.....	67
Lista de símbolos, abreviaturas y siglas.....	74
Lista de tablas, figuras y gráficos.....	75
Bibliografía.....	77
Apéndice 1.....	86
Apéndice 2.....	95
Apéndice 3.....	104

1.- PREDICCIONES DE RECONOCIMIENTO EPISÓDICO.

¿Podemos predecir que reconoceremos en el futuro a una persona que nos acaban de presentar? ¿Podemos predecir qué ítems responderemos correctamente en una prueba de reconocimiento? Las pruebas de reconocimiento de memoria son utilizadas habitualmente en educación para verificar la información aprendida por los estudiantes. Cuando un alumno de bachillerato prepara un examen de biología sobre los nematelmintos, le surgen multitud de pensamientos mientras estudia: “no lo comprendo”, “este tema es muy difícil”, “seguro que no reconoceré estas palabras tan raras en el examen”. Pues bien, hablamos de metacognición cuando nuestros pensamientos se refieren a la propia actividad cognitiva, es decir, al conocimiento de nuestro propio conocimiento. Cuando el alumno de bachillerato afirma “no lo comprendo”, está emitiendo un juicio acerca del estado de su propio conocimiento sobre la materia estudiada. Además, estos juicios le permiten anticipar mentalmente su futuro personal: “seguro que no reconoceré estas palabras tan raras en el examen”.

Los procesos metacognitivos, que regulan y controlan varios aspectos del procesamiento de la información y de la conducta, han sido objeto de aumento de interés y de estudio en las últimas décadas (Flavell, 1979; Nelson y Narens, 1990; Metcalfe y Shimamura, 1994; Perfect y Schwartz, 2002; Koriat, 2012; Dunlosky y Thiede, 2013). Un apartado importante en el estudio de esta área es la metamemoria, que hace referencia al conocimiento y conciencia que tenemos de nuestra memoria y a cómo este conocimiento influye en la supervisión y control del propio aprendizaje y la recuperación de la información.

De acuerdo con el modelo propuesto por Nelson y Narens (1990; ver figura 1.1 a continuación), los dos aspectos fundamentales de los procesos metacognitivos

son la *supervisión* y el *control*. Por ejemplo, un estudiante que prepara una determinada asignatura, supervisa su grado de conocimiento de la materia mediante sus autoevaluaciones o juicios sobre su nivel de aprendizaje, controlando a su vez su conducta a partir de ahí, para dedicar o no un mayor tiempo de estudio.

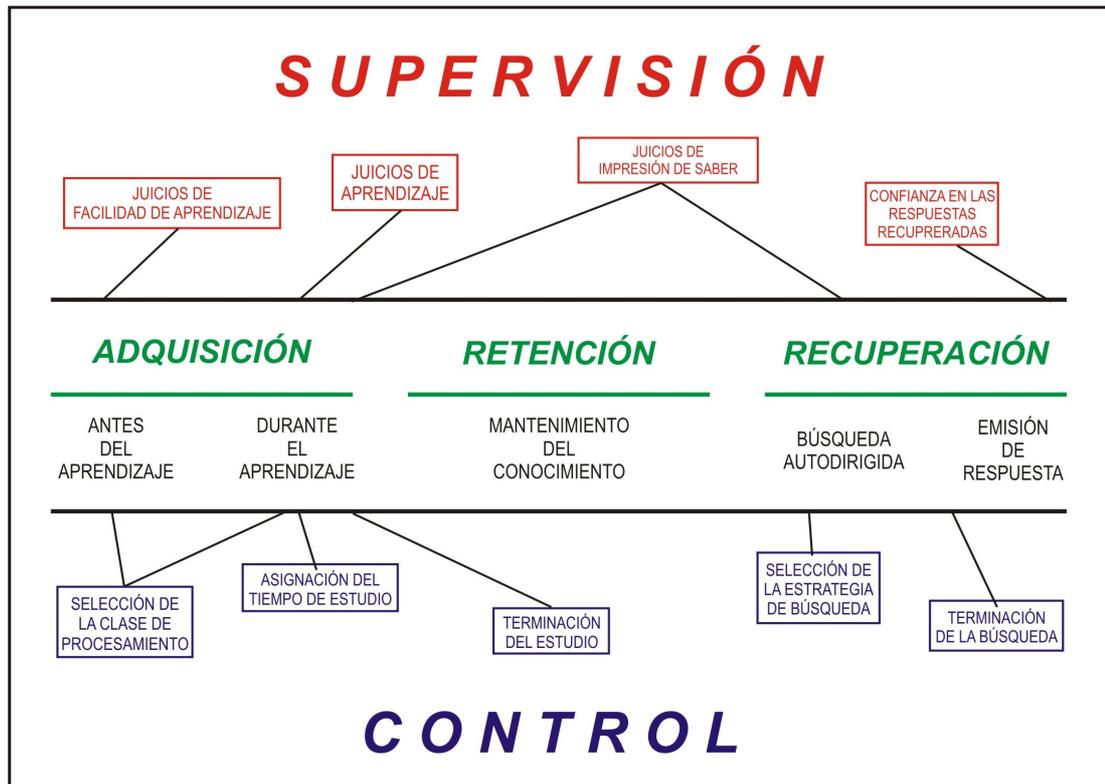


Figura 1.1. Adaptación del esquema de metamemoria propuesto por Nelson y Narens (1990). En el centro aparecen las etapas del proceso de memoria, en la parte superior algunos ejemplos de componentes de supervisión y en la parte inferior algunos ejemplos de componentes de control.

Los procedimientos básicos de investigación en este campo pueden ser agrupados dependiendo del momento en que se realizan los juicios metamnemónicos. Nelson y Narens (1990) propusieron cuatro categorías de juicios en función del momento en que se producen: (1) *Juicios de Facilidad de aprendizaje* (JFA); son los juicios que se realizan antes de la adquisición, sobre los ítems que van a ser aprendidos y son estimaciones sobre su facilidad o dificultad. (2) *Juicios de Aprendizaje* (JA); son los que se realizan durante o después de la adquisición, y son predicciones sobre el rendimiento en una futura prueba de memoria. (3) *Juicios de*

Impresión de Saber (JIS); ocurren durante la recuperación de la información, son juicios acerca de si los ítems no recordados en el momento de la recuperación podrán ser recuperados o no en una futura prueba de memoria. (4) *Juicios de Confianza en las respuestas recuperadas* (JC); ocurren después la recuperación de la información, las personas tienen que valorar su confianza en las respuestas emitidas. Las tres primeras pertenecen a la supervisión prospectiva y la cuarta a la supervisión retrospectiva.

En esta investigación vamos a trabajar con el paradigma experimental de *Juicios de Aprendizaje* (JA) en el ámbito del aprendizaje clásico de pares asociados. El procedimiento de esta técnica es el siguiente: en primer lugar se realiza una tarea de aprendizaje de pares asociados de palabras, en la que las personas estudian el par estímulo-respuesta (p. ej. guerra-número). Después realizan juicios de aprendizaje en los que se pregunta a los participantes: *¿Cuál es la probabilidad de recuperar la palabra respuesta del par cuando se le presente la palabra estímulo?* Para finalizar, se realiza la prueba de memoria donde el participante debe recuperar la palabra respuesta ante la presentación de la palabra estímulo. El trabajo de Arbuckle y Cuddy (1969) es pionero en este campo. Para una revisión véase: Zechmeister y Niberg (1982); Leonasio y Nelson (1990) Dunlosky y Nelson (1994) y Ruiz (2004).

2.- APROXIMACIONES TEÓRICAS EN METAMEMORIA.

La *hipótesis del acceso directo al trazo* es la teoría clásica en la que, de un modo más o menos explícito, la mayoría de las personas creen. Según esta teoría, las personas tienen un mecanismo que supervisa el trazo de memoria en cada momento y pueden así conocer el estado o calidad de su contenido. En el ámbito más restringido

de la psicología experimental fue propuesta en primer lugar por Hart (1965, 1967) para los *Juicios de Impresión de Saber* (JIS), quien denominó a dicho mecanismo MEMO, del Inglés *Memory Monitoring*. Según Hart, al basar MEMO sus juicios en el contenido mismo de la memoria, se deduce que los resultados de ambos mecanismos están asociados; es decir, debe haber una asociación funcional entre los juicios metacognitivos y el rendimiento memorístico. Esta hipótesis tiene, por tanto, dos implicaciones fundamentales: (a) cualquier factor o variable que aumente o disminuya el rendimiento de memoria influirá en el mismo sentido sobre los juicios metamnemónicos; y (b) las predicciones de las personas deben ser acertadas, aunque admitamos cierto margen de error.

Algunos resultados experimentales han apoyado esta hipótesis (p. ej., Nelson, Gerler y Narens, 1984; Burke, Mackay, Worthley y Wade, 1991; James y Burke, 2000). Por ejemplo, el trabajo de James y Burke (2000) reveló que el procesamiento de palabras relacionadas fonológicamente reducía la aparición del fenómeno de *sensación de la punta de la lengua*¹ e incrementaba la recuperación de respuestas a cuestiones de conocimiento general. Estos resultados defienden el modelo de *déficit de transmisión*, el cual, sostiene que el fenómeno de la *sensación de la punta de la lengua* es causado por conexiones débiles entre las representaciones fonológicas.

¹ La *sensación de la punta de la lengua* es un fenómeno popular que todos hemos experimentado alguna vez (vemos un artista en televisión y decimos: "tengo su nombre en la punta de la lengua", tenemos la certeza de conocer su nombre aunque en ese momento no podemos recordarlo). Brown y McNeill (1966) fueron los primeros en estudiar este fenómeno experimentalmente. A nivel de procedimiento, se diferencia de los *juicios de impresión de saber* (JIS) en que este fenómeno indica la sensación subjetiva de que el recuerdo es inminente, mientras que, los JIS evalúan si las personas creen que reconocerán el ítem no recordado (Schwartz, 1994).

2.1.- DISOCIACIÓN ENTRE LAS PREDICCIONES Y EL RENDIMIENTO MEMORÍSTICO.

En general, hablamos de disociación cuando una variable tiene efectos diferenciales sobre dos o más medidas de memoria. En nuestro ámbito, podemos ver un ejemplo de disociación funcional en el experimento 1 de Begg, Duft, Lalonde, Melnick, y Sanvito (1989). En este caso, entre otras manipulaciones, estos investigadores variaron la frecuencia de las palabras. Los participantes estudiaban una lista de palabras al tiempo que estimaban su memorabilidad posterior (juicios de aprendizaje) en una escala de 7 puntos. En la parte final del procedimiento realizaban una prueba de reconocimiento. Se sabe como es el efecto de la frecuencia de las palabras sobre el reconocimiento: las palabras de baja frecuencia (raras) se reconocen mejor que las palabras de alta frecuencia (comunes). El efecto de la frecuencia fue formulado en primer lugar por Hall (1954) y Sumbly (1963) en pruebas de recuerdo y por Gorman (1961) en pruebas de reconocimiento (véase p. ej., Gregg, 1976; Schulman, 1976; Malmberg y Nelson, 2003; Coane, Balota, Dolan y Jacoby, 2011). En el experimento 1 de Begg y otros los participantes atribuyeron en sus predicciones una mayor memorabilidad a las palabras comunes que a las raras; por el contrario, en la prueba de reconocimiento las palabras raras tuvieron un mayor número de aciertos que las comunes. Tenemos por tanto, que la variable independiente frecuencia de las palabras produce efectos contrapuestos en las dos medidas: predicciones y prueba de reconocimiento (variables dependientes).

La investigación ha demostrado que hay otros factores que también producen disociaciones claras entre los juicios metamnemónicos y el rendimiento memorístico. Entre ellos, podemos citar cuál sea la tarea de estudio (Dunlosky y Nelson, 1994), la composición de la lista de estudio (Susser, Mulligan y Besken, 2013), la repetición

del ítem en la lista de estudio y su tiempo de presentación (Koriat, 1997), el hecho de caminar antes del estudio (Salas, Minakata y Kelemen, 2011), o las expresiones emocionales faciales (Nomi, Rhodes, y Cleary, 2013).

Lo importante para nosotros de tales disociaciones es que, según el mecanismo MEMO de la *hipótesis del acceso directo al trazo*, debe haber una *asociación funcional* entre las predicciones y el rendimiento de memoria; por consiguiente, esta hipótesis no puede dar una explicación de las disociaciones. La cuestión ahora es explicar por qué, si no hay asociación funcional entre predicciones y rendimiento, generalmente los juicios metamnemónicos suelen ser al menos aceptablemente válidos. Pues bien, la alternativa propuesta por los investigadores para explicar la precisión de los juicios metacognitivos es la denominada *hipótesis de los mecanismos inferenciales* (Nelson y Narens, 1990, 1994; Koriat, 1997), aunque no siempre aparece bajo esta denominación (p. ej., la hipótesis de la "utilización de indicios" de Koriat, 1997).

2.2.- FUENTES DE INFORMACIÓN EN METAMEMORIA.

Para la hipótesis inferencial los juicios de metamemoria se elaboran mediante *heurísticos* que se basan en diferentes fuentes de información. Un heurístico puede ser considerado como una regla o estrategia simple que utilizamos para emitir los juicios. Si te preguntan si un compañero de trabajo es simpático y el primer recuerdo que te viene a la mente es que el otro día estaba contando un chiste en el bar, puede que digas que es simpático; pero si el recuerdo más accesible es verle enfadado con otro compañero, dirás que es antipático. En muchas ocasiones este heurístico puede resultar útil, puesto que si acceden a nuestra mente muchos ejemplos de esta persona

siendo simpática, probablemente sea debido a que en la realidad lo es (Kahneman, 2012; Kahneman y Frederick, 2005; Kahneman y Tversky, 1973; Tversky y Kahneman, 1974).

En el ámbito concreto de la metamemoria las principales claves mnemónicas que se han propuesto como fuentes de información sobre las que elaborar las inferencias metacognitivas pueden agruparse en las siguientes categorías: recuperabilidad del objetivo (Dunlosky y Nelson, 1992, 1994, 1997; Nelson y Dunlosky, 1991; Nelson, Narens y Dunlosky, 2004), accesibilidad de información de algún modo relacionada con el contenido cuyo recuerdo o reconocimiento se predice (Koriat, 1993, 1994; Koriat y Levy-Sadot, 2001), familiaridad de la clave o estímulo que actúa como contexto del juicio (Metcalf, 1993; Metcalf y Finn, 2008; Metcalf, Schwartz y Joaquim, 1993; Reder y Ritter, 1992; Schwartz y Metcalf, 1992; Son y Metcalf, 2005) y facilidad de procesamiento de ese mismo contexto o de la tarea de predicción (Begg y otros, 1989; Hertwig, Herzog, Schooler y Reimer, 2008; Undorf y Erdfelder, 2011). Veamos qué hay tras estas hipótesis con algo más de detalle.

Cuando se utilizan *juicios demorados*² todo indica que la estrategia más habitual puede ser la evaluación tácita de la recuperabilidad del objetivo. En el momento en que las personas toman sus juicios de aprendizaje, supervisan la información recuperada desde la memoria acerca del ítem que va a ser juzgado; es la llamada *Hipótesis de la Supervisión de la Recuperación* (Dunlosky y Nelson, 1994, 1997; Nelson, Narens y Dunlosky, 2004). Los juicios metacognitivos están

² El *efecto de la demora del juicio de aprendizaje* (Dunlosky y Nelson, 1992 ; Nelson y Dunlosky, 1991) nos dice que la precisión de las estimaciones es mayor cuando los juicios de aprendizaje se realizan un corto tiempo después del estudio (juicios demorados) que cuando se realizan inmediatamente después del mismo (juicios inmediatos). Es un fenómeno robusto que se da en múltiples circunstancias: distintos métodos de estudio (Dunlosky y Nelson, 1994), en pruebas de reconocimiento (Thiede y Dunlosky, 1994; Dunlosky y Nelson, 1997) o en personas mayores (Connor, Dunlosky y Hertzog, 1997). Para una revisión véase Rhodes y Tauber (2011).

determinados por la recuperabilidad de la información. Por lo tanto, esta hipótesis pronostica una asociación funcional entre el rendimiento de memoria y los juicios: cuando una variable produce un rendimiento memorístico elevado, provocará a su vez, que las estimaciones de memoria sean también elevadas.

Sin embargo, hoy sabemos que en el procedimiento RJR (recuerdo-juicio-reconocimiento) de los *Juicios de Impresión de Saber* las estimaciones de las personas se realizan sobre los ítems que no han sido recordados, siendo su precisión relativamente aceptable (Hart, 1965, 1967; Schacter, 1983). Lo cual, nos indica que la recuperabilidad no agota la explicación de los juicios metamnemónicos (Schwartz, 1994).

Por su parte, el heurístico de accesibilidad propuesto por Koriat (1993, 1994) supone que los *Juicios de Impresión de Saber* (JIS) están basados en cualquier tipo de información recuperada acerca del objetivo (sea correcta o incorrecta) y en la intensidad de la información parcial recuperada (por ejemplo: si la información parcial es particularmente vívida, específica o persistente). Obsérvese que se diferencia este heurístico del de la recuperabilidad del objetivo, en que no limita la información parcial a la información correcta, sino que cualquier tipo de información puede ser usada en los juicios.

Por otra parte, algunos estudios han demostrado que las personas pueden basar sus estimaciones en el grado de familiaridad asociado con el propio contexto de recuperación (Reder y Ritter, 1992; Schwartz y Metcalfe, 1992). Se trata de información no necesariamente relacionada con el objetivo que está presente en el momento de los juicios y que es utilizada como clave para las predicciones (Metcalfe y otros., 1993; Son y Metcalfe, 2005).

Para Begg y otros (1989) las predicciones de memoria están basadas en la facilidad de procesamiento de los ítems durante la propia actividad de elaboración del juicio metamnemónico. Se trata del heurístico implícito *lo que es fácil de hacer ahora, será recordado mejor luego* (Begg y otros, 1989, pág. 610). El heurístico de facilidad de procesamiento nos dice que la precisión de las estimaciones se debe a motivos puramente correlacionales, ya que muchas variables que producen un incremento de la memorabilidad también producen un incremento de la facilidad de procesamiento. Esta covariación hace razonable el uso de la facilidad de procesamiento como base para las estimaciones de memoria (véase, por ejemplo, Undorf y Erdfelder, 2011).

Para finalizar, debemos añadir que las fuentes de información se pueden combinar dinámicamente para formar los heurísticos sobre los que se van a basar las estimaciones, tanto en las fases de adquisición y retención, como en la recuperación de la información, donde el contexto estimular y la experiencia personal también juegan un papel importante (Koriat, 1997). Por ejemplo, Son y Metcalfe (2005; también Metcalfe y Finn, 2008) han estudiado cómo se combinan la familiaridad y la recuperabilidad (aunque véase una propuesta alternativa en Hertwig, Herzog, Schooler y Reimer, 2008); Koriat y Ma'ayan (2005) han investigado la contribución conjunta de la fluidez de codificación y la de recuperación; y finalmente, Koriat y Levy-Sadot (2001) la combinación de la familiaridad de la clave y la accesibilidad.

Hasta aquí, hemos repasado brevemente diferentes proposiciones sobre las fuentes de información que la persona utiliza para elaborar sus juicios metacognitivos. Pero hemos dejado para el final la propuesta sobre la que tratará este trabajo. Se trata de la denominada hipótesis de la *Supervisión Apropiaada para la Transferencia* como explicación de la precisión o imprecisión de los juicios metacognitivos. Aunque

antes de analizar esta hipótesis, debemos examinar sus antecedentes teóricos: el *Procesamiento Apropiado para la Transferencia*.

3.- PROCESAMIENTO APROPIADO PARA LA TRANSFERENCIA.

De acuerdo con el *principio de codificación específica*, el modo en que se codifica un evento determina los modos de recuperación que serán eficaces para poder acceder a lo almacenado (Tulving y Thomson, 1973).

En línea con lo anterior, Morris, Bransford y Franks (1977) realizaron una serie de experimentos para someter a prueba hipótesis derivadas del enfoque de los niveles de procesamiento (Craik y Lockhart, 1972; véase también Craik, 2002; Lockhart, 2002; Rose y Craik, 2012; Rose, Myerson, Roediger y Hale, 2010; Ruiz-Vargas y Cuevas, 1999). Sus hallazgos mostraron que la relación entre la codificación inicial de la información y su recuperación posterior es fundamental en los procesos de memoria. Estos autores formularon la hipótesis del *Procesamiento Apropiado para la Transferencia* (del Inglés *Transfer-appropriate processing*), según la cual el rendimiento de la memoria es mejor cuando el *procesamiento* utilizado durante la recuperación es similar al *procesamiento* utilizado durante la codificación de la misma información recuperada.

Desde el artículo de Morris y otros (1977), la perspectiva del *Procesamiento Apropiado para la Transferencia* ha generado un gran número de investigaciones que llegan hasta nuestros días. Con hallazgos tanto a su favor (véase p. ej., Bugg y McDaniel, 2012; Parks, 2013; Park y Rugg, 2008) como en su contra (véase p. ej., McBride y Abney, 2012; Poirier, Nairne, Morin y Zimmermann, 2012).

4.- SUPERVISIÓN APROPIADA PARA LA TRANSFERENCIA.

La *Supervisión Apropiaada para la Transferencia* (SAT; en Inglés *Transfer-appropriate monitoring*) es una extensión de la hipótesis del *Procesamiento Apropiaado para la Transferencia*. La hipótesis SAT nos dice que la precisión de los juicios metamnemónicos estará en función de la similitud entre las condiciones durante los juicios y las condiciones durante las pruebas de memoria. A mayor similitud entre las condiciones de juicios metamnemónicos y pruebas de memoria, mayor similitud entre los procesamientos ligados a unos y otras y, en consecuencia, mayor precisión de los juicios metamnemónicos (Begg y otros, 1989; Dunlosky y Nelson, 1997; Dunlosky, Rawson y Middleton, 2005; Kennedy y Nawrocki, 2003; Koriat, 1993, 1994; Weaver y Kelemen, 2003).

Por ejemplo, en el experimento 4 de Begg y otros (1989) un grupo de universitarios estudiaron pares asociados de palabras AB. Después de emitir los juicios de aprendizaje, recibieron una prueba de memoria final de reconocimiento de A y otra de recuerdo de A→B (se presentaba el término estímulo A y debían recordar el término respuesta B). El objetivo de este experimento era contrastar el efecto de dos variables presentes en las condiciones de emisión de los juicios: las claves de estimación (factor contextual) y el tipo de predicción (factor estratégico). Las claves de estimación fueron de tres tipos: el término estímulo del par (A), el término respuesta (B), o ambos (AB). El tipo de predicción dependía de la pregunta que se les planteaba a los participantes: se les pedía que realizaran predicciones de reconocimiento de A, predicciones de reconocimiento de B, o predicciones de recuerdo del término respuesta en presencia del término estímulo (A→B). Los resultados mostraron que las predicciones eran más ajustadas a las pruebas de reconocimiento cuando el ítem mostrado durante los juicios era el mismo que iba a

ser examinado en reconocimiento (A). Por otro lado, las predicciones eran más ajustadas al recuerdo cuando se cumplían dos condiciones: (a) el estímulo para los juicios era el término estímulo del par, con la palabra objetivo ausente; y (b) las instrucciones de la tarea de juicio hacían referencia al estímulo ausente. Un aspecto crucial de tales resultados es que ambos se producían independientemente de la predicción explícitamente pedida a los participantes. Un ejemplo paradigmático de esto es que la predicción del reconocimiento de B cuando ésta se hace ante A, está tan bien calibrada respecto de la prueba de recuerdo como la propia predicción del recuerdo de B. Algo similar encontramos en los ajustes de los juicios al reconocimiento de A: es igual cuando se pide predicción de recuerdo (o también reconocimiento) de B que cuando se pide predicción de reconocimiento de A. Paradójicamente, cuando los participantes predicen reconocimiento de B ante A, los juicios son más ajustados al reconocimiento de A que cuando predicen el propio reconocimiento de A en el contexto de B.

Thiede y Dunlosky (1994) en su experimento 2 obtuvieron resultados análogos. Durante la fase de estudio y juicios de aprendizaje, un grupo de participantes era instruido para anticipar una prueba de recuerdo de pares asociados y otro grupo era instruido para anticipar una prueba de reconocimiento de elección forzosa entre diez alternativas. Ambos grupos recibieron las dos pruebas de memoria. Los resultados mostraron diferencias significativas para el tipo de prueba recibida, pero no mostraron diferencias significativas para el tipo de prueba anticipada.

Por tanto, parece que son las condiciones estimulares de la tarea (factor contextual) las que predominan al determinar el tipo de procesamiento que utilizarán las personas, con independencia de la prueba de memoria que esperen (factor estratégico). Y la calibración de los juicios metamnemónicos basados en ese

procesamiento provocado por el estímulo parece depender, a su vez, de su similitud con el procesamiento necesario para la prueba final de memoria. De ahí que el ajuste al reconocimiento requiera sólo la presencia aislada en el juicio del estímulo de reconocimiento, mientras que el ajuste al recuerdo exija, además, la referencia al término B en la misma tarea de juicio, cualquiera que ésta sea.

5.- CRÍTICAS A LA HIPÓTESIS SAT.

Pero no siempre los resultados obtenidos por los investigadores han respaldado la hipótesis de la SAT. Así, Dunlosky y Nelson (1997) realizaron tres experimentos de aprendizaje de pares asociados, donde los participantes realizaban juicios de aprendizaje demorados utilizando como clave el estímulo solo o el par estímulo-respuesta. Posteriormente realizaban una prueba de memoria final de reconocimiento asociativo³. Nótese que el contexto estimular de esta prueba de memoria es más similar al de los juicios de aprendizaje emitidos ante el par estímulo-respuesta que al de los juicios de aprendizaje emitidos ante el estímulo solo. Por este motivo, la SAT pronostica que los juicios de aprendizaje serán más precisos cuando utilicemos como clave el par estímulo-respuesta que cuando utilicemos como clave el estímulo solo. Sin embargo, los resultados mostraron que la precisión de los juicios ante el estímulo solo era superior que los juicios ante el par estímulo-respuesta, contradiciendo los pronósticos de la SAT.

Pocos años después, Weaver y Kelemen (2003) hicieron una investigación

³ Reconocimiento asociativo implica el reconocimiento por parte de la persona de que la palabra estímulo y respuesta han sido estudiados juntos. La persona debe seleccionar el par estudiado entre otros pares reordenados (distractores) de la lista de estudio. Mientras que, en prueba de reconocimiento del ítem se presenta una sola palabra a los participantes y estos tienen que decir si esta palabra era antigua o nueva, antigua cuando pertenece a la lista de estudio y nueva cuando pertenece a dicha lista.

más detallada de la SAT, que también arrojó resultados contrarios a ella. En este trabajo los participantes estudiaron pares asociados de palabras (p. ej.: ELEFANTE-bronceado). Después emitieron juicios de aprendizaje para predecir su ejecución en una prueba de recuerdo o reconocimiento bajo 5 condiciones: en la condición 1 la clave era el estímulo solo (p. ej.: "ELEFANTE- ¿?"); en la condición 2 la clave era el par estímulo-respuesta (p. ej.: "ELEFANTE-bronceado"); en la condición 3 la clave era el estímulo solo y otras seis alternativas consistentes en el estímulo unido a respuestas incorrectas (p. ej.: "ELEFANTE-diamante", "ELEFANTE- ¿?", "ELEFANTE-azúcar",...); en la condición 4 la clave era como la anterior, pero substituyendo el estímulo solo por el par estímulo-respuesta correcta (p. ej.: "ELEFANTE-diamante", "ELEFANTE- bronceado", "ELEFANTE-azúcar",...) y en la condición 5 la clave era el par estímulo respuesta marcado con tres asteriscos más seis alternativas estímulo+respuesta incorrectas (p. ej.: "ELEFANTE-diamante", "ELEFANTE- bronceado***", "ELEFANTE-azúcar", ...). Para finalizar, un grupo de participantes recibió una prueba de recuerdo con clave y otro grupo recibió una prueba de reconocimiento asociativo de elección forzosa entre siete alternativas.

Pues bien, según la hipótesis SAT, respecto a la prueba de recuerdo debemos esperar una mejor precisión de las estimaciones en la condición 1 (en adelante C1) que en las condiciones 2 (C2), 4 (C4) y 5 (C5), porque en esta condición se establece una correspondencia exacta tanto en contexto como presumiblemente también en procesamiento entre juicios y prueba de recuerdo con claves. Mientras que, respecto a la prueba de reconocimiento se esperan juicios más acertados en la condición 4 (C4) que en las condiciones 1 (C1) y 3 (C3), porque en esta condición se supone una mayor correspondencia contextual y estratégica entre juicios y reconocimiento asociativo. Experimentalmente, se esperaba, por tanto, la interacción entre juicios y

pruebas de memoria. Pese a ello, esta interacción no se produjo. Las predicciones de la SAT se cumplen en relación con las estimaciones de recuerdo pero no se cumplen en relación con las estimaciones de reconocimiento.

Por lo tanto, estos resultados van en contra de los pronósticos de la SAT. Una vez más, en general la precisión de las estimaciones era superior cuando el término respuesta no estaba presente en el momento de los juicios. Weaver y Kelemen (2003) nos dicen que estos hallazgos invitan a pensar que los participantes intentaban recuperar la respuesta correcta durante los juicios en las condiciones 1 y 3, siendo el éxito o no del intento un criterio para hacer la predicción. Por el contrario, al no ser necesaria la recuperación en las condiciones 2 y 5, éstas ofrecen pocas claves diagnósticas para los juicios, ya que está presente el par estímulo-respuesta durante las predicciones (Begg y otros, 1989; Dunlosky y Nelson, 1994, 1997). Hay que señalar la similitud de las condiciones 1 y 3 de Weaver y Kelemen (2003), por una parte, y 2 y 5, por otra, con las condiciones análogas de Begg y otros, (1989, experimento 4) y Dunlosky y Nelson (1997) que se han descrito arriba.

6.- PLANTEAMIENTO DEL PROBLEMA Y FORMULACIÓN DE HIPÓTESIS.

En el trabajo de Weaver y Kelemen (2003) no se encuentran pruebas a favor de la SAT. Según ellos, habiendo correspondencia exacta, tanto en contexto como en procesamiento, entre juicios y reconocimiento (RN) en C4, no se explica que C4 no sea más predictiva del RN⁴ que C1, donde el contexto es más similar a recuerdo (RD). Por tanto, la pregunta ahora es: ¿por qué C4 no es más predictiva de RN que

⁴ En Weaver y Kelemen (2003) pág, 1061 se lee: "Condition 4 provided an exact match of context and processing at JOL and at test".

C1, existiendo una correspondencia aparentemente exacta entre contextos y procesamientos?

Podemos ver una pista para dar con la solución en un dato de este trabajo. Estos autores pasan por alto en sus resultados que C4 no sólo es tan predictiva de reconocimiento como C1, sino que además, también es tan predictiva de recuerdo (RD) como C1, a pesar de que en C4 no existe correspondencia entre los contextos de juicio y de prueba.

Nosotros pensamos que, el hecho de que C4 sea predictiva de RD se explicaría si asumimos que en los juicios de C4 los participantes activan procesos de recuperación análogos a los que ponen en marcha en RD. Generalmente se admite que una información necesaria para el recuerdo con clave (pares asociados) es la asociación entre la clave y el objetivo (*información asociativa*), por lo que nosotros suponemos que ésta sería la información utilizada por los participantes en los juicios de C4. Sólo así los juicios en C4 podrían ser tan buenos predictores del rendimiento en la prueba de recuerdo de Weaver y Kelemen (2003) como los juicios en C1.

En consonancia con nuestra hipótesis está el hecho de que las condiciones contextuales de C4 se asemejan a las de una prueba típica de reconocimiento asociativo.

Desde la segunda mitad del siglo pasado, muchos estudios han sugerido la intervención de dos procesos en reconocimiento: recuperación y familiaridad (p. ej., Atkinson y Joula, 1974; Jacoby y Dallas, 1981; Mandler, 1980, 1991; véase una revisión en Ruiz, 2004). Una prueba de reconocimiento implica un complejo juego de procesos donde se combinan la recuperación y la familiaridad⁵. La persona puede

⁵ Tulving (1985) denomina a estos dos procesos: "recordar" y "saber". Introduce un nuevo enfoque metodológico para evaluar la experiencia subjetiva de reconocimiento. Después de estudiar pares de palabras AB, cuando el reconocimiento del ítem B va acompañado del recuerdo consciente de que ese

basar sus respuestas en recuperaciones concretas del contexto de adquisición, lo cual, le permite asegurar que recuerda la palabra objetivo. O simplemente experimenta una sensación subjetiva de familiaridad con el objetivo, sensación que en ocasiones puede ser suficiente para emitir un reconocimiento positivo. En otras palabras, se supone que en una prueba de reconocimiento episódico la familiaridad actúa en combinación con la recuperación de la información para hacer que la persona reconozca o no la palabra objetivo.

Pues bien, existen pruebas en favor de la hipótesis de que en condiciones de *reconocimiento asociativo* los individuos hacen uso de ambos tipos de informaciones. La primera estaría ligada al componente de recuperación en el reconocimiento y la segunda al componente de familiaridad. Veamos algunos ejemplos.

Unos resultados de Clark (1992) mostraron que el efecto de la frecuencia de las palabras se da en reconocimiento del ítem (RI) pero no se da en reconocimiento asociativo (RA). De acuerdo con este autor, estos resultados sugieren que en reconocimiento asociativo tienen un peso considerable los procesos de recuperación, como ocurre en recuerdo. En línea con este argumento, Clark y Shiffrin (1992) obtuvieron un efecto de la frecuencia intermedio entre recuerdo y reconocimiento con un procedimiento que denominaron *cued recognition*⁶.

Por su parte, Malmberg, Zeelenberg y Shiffrin (2004) concluyeron que la familiaridad es la mejor estrategia cuando los distractores son *diferentes*

ítem B había sido estudiado junto con A, tenemos una experiencia del tipo "recordar". Cuando el reconocimiento del ítem B va acompañado exclusivamente de una sensación de familiaridad con él, pero sin recordar nada del contexto de adquisición, tenemos una experiencia del tipo "saber".

⁶ La expresión *cued recognition* se refiere a la situación en que la persona ve en la prueba de reconocimiento pares AB (antiguo) o AX (nuevo), sabiendo de antemano que del par de prueba el primero es una palabra de la lista de pares, mientras que el segundo puede que sí o puede que no lo sea. Es una situación análoga a C2 en Weaver y Kelemen (2003) y Dunlosky y Nelson (1997).

*aleatoriamente*⁷. De modo complementario, Malmberg, Holden y Shiffrin (2004), hallaron que la recuperación es más eficiente que la familiaridad cuando hay una gran similitud entre objetivos y distractores, como es el caso del reconocimiento del número gramatical (o discriminación de la pluralidad, como *mono* vs. *monos*). Xu y Malmberg (2007) muestran cómo el modelo de doble proceso de Malmberg, Zeelenberg y Shiffrin (2004) y Malmberg, Holden y Shiffrin (2004; véase también Malmberg, 2008) puede aplicarse también para explicar el efecto de la repetición y el tipo de estímulo sobre aciertos y, muy en especial, sobre las falsas alarmas en un procedimiento de RA (véase también Jang, Pashler y Huber, 2014).

Otros resultados de estos investigadores también son congruentes con lo anterior. Así Malmberg y Xu (2007, experimento 3) demuestran en RA que con la mayor disponibilidad del tiempo de respuesta de reconocimiento los participantes hacen recaer más el peso de su decisión de rechazo de distractores en información asociativa. Mientras que con la introducción de pares XY⁸ (su experimento 4) entre los distractores, la tasa general de falsas alarmas crece, pero la tasa específica ligada a estos estímulos (XY) es relativamente muy pequeña. Lo cual, sugiere, que hay más intervención del componente de familiaridad en general (de ahí el aumento global de falsas alarmas) y que gracias a él resulta fácil descartar los pares XY presentes en la lista (debido a su baja familiaridad relativa). En general, este grupo de investigadores concluyen de sus resultados que, como corresponde a la naturaleza flexible y falible de la memoria humana, las personas utilizan diferentes estrategias de reconocimiento asociativo inducidas por factores situacionales (Malmberg, 2008).

El estudio de los efectos clásicos de la interferencia en pruebas de pares

⁷ Ellos usan la expresión "randomly different". Donde "random" se refiere a que no se controla el atributo o propiedad por la que el objetivo se diferencia de los distractores.

⁸ Se trata de pares asociados compuestos por términos estímulo y respuesta que no han aparecido en la fase de estudio, por lo que se supone que suscitan un grado mínimo de familiaridad.

asociados también ha aportado información sobre la implicación de dos procesos en tareas de emparejamiento asociativo. Así, por ejemplo, Verde (2004) ha mostrado que ciertas condiciones clásicas de interferencia (A-B, A-D) reducen los procesos de recuperación pero incrementan el peso de la familiaridad. En línea con esto, Buchler, Light y Reder (2008) mostraron también cómo el fenómeno de interferencia conocido como *efecto abanico* aparece específicamente en RA; y cómo los individuos son capaces de distinguir explícitamente entre ítems que han aparecido emparejados y aquellos otros que también han aparecido, pero desemparejados. Sus datos en conjunto apoyan la idea de que la información asociativa y del ítem activan diferentes formas de procesamiento y contribuyen de forma separada a las tareas de recuperación.

Interesan especialmente aquí unos datos obtenidos por Yonelinas (1997) al poner a prueba su hipótesis de que el componente de familiaridad en una tarea de reconocimiento funciona de acuerdo con un modelo de *Teoría Detección de Señales* (véase, por ejemplo, Green y Swets, 1966; Ruiz, 2003), mientras que el componente de recuperación funciona según los supuestos de la *Teoría Psicofísica del Umbral Alto* (Luce, 1963). Yonelinas pudo demostrar cómo, en condiciones de emparejamiento asociativo en las que se neutraliza el papel de la familiaridad, la función COR⁹ es lineal. Esta linealidad, predicha sólo por la Teoría del Umbral Alto, sería consecuencia, según Yonelinas, del predominio del componente de recuperación en sus condiciones de reconocimiento asociativo, si asumimos con él que este componente funciona de acuerdo con un modelo de decisión de umbral alto (véase una revisión en Ruiz, 2004).

⁹ Las curvas COR (Característica Operativa del Receptor) son las funciones que relacionan las tasas de aciertos y falsas alarmas. Esas funciones, en coordenadas lineales, presentan una forma *curva* típica en el ámbito de la *Teoría de Detección de Señales*. Y una forma *lineal* típica en el ámbito de la *Teoría del Umbral Alto*.

También hay estudios evolutivos que apoyan la hipótesis del procesamiento dual en reconocimiento. Por ejemplo, Castel y Craik (2003) estudiaron cómo influye el envejecimiento en la información asociativa y del ítem. Encuentran que el RA presenta un deterioro con la edad que no va acompañado del mismo deterioro en reconocimiento de ítem. En esta misma línea Naveh-Benjamin (2000, especialmente experimento 2) observaron que las personas mayores mostraron un mayor déficit asociativo de memoria que los jóvenes. (véase también: Li, Naveh-Benjamin, y Lindenberger, 2005; Old y Naveh-Benjamin, 2008; Ratcliff, Thapar y McKoon, 2011).

Además de estas investigaciones genuinamente comportamentales, algunos autores han dado un paso más, sugiriendo que recuerdo y familiaridad son procesos que se apoyan en regiones del cerebro parcialmente independientes (p. ej., Haskins, Yonelinas, Quamme, y Ranganath, 2008; Sauvage, Fortin, Owens, Yonelinas y Eichenbaum, 2008; Yonelinas, Aly, Wang y Koen, 2010; puede verse una revisión en Lee, Fell y Axmacher, 2013).

Todos estos resultados respaldan en cierto modo la hipótesis que adelantamos páginas atrás, para explicar el ajuste al rendimiento en recuerdo de las predicciones hechas en la condición C4 de Weaver y Kelemen (2003): una intervención considerable de la información *asociativa* en la tarea misma de predicción. Ahora bien, si los juicios de C4 tienen un fuerte componente de información asociativa por la composición del contexto, también lo debe tener la prueba de reconocimiento de Weaver y Kelemen (2003). En tal caso, ¿por qué no hay más precisión de los juicios respecto al reconocimiento en C4 que respecto al recuerdo?

Esta inferior precisión de las estimaciones ante pruebas de reconocimiento con respecto a pruebas de recuerdo fue investigada por Thiede y Dunlosky (1994).

Estos autores sugieren que este hecho está provocado por algún factor que ocurre durante la prueba de reconocimiento. Entre los posibles factores señalan la eliminación de alternativas incorrectas y las adivinaciones correctas.

Pero podemos encontrar la respuesta a esta pregunta también en los modelos de procesos duales del reconocimiento: el juego combinado de familiaridad y recuperación. Malmberg y Xu (2007) afirman que los individuos no son capaces de controlar ensayo a ensayo el peso de la recuperación y la familiaridad, pero a lo largo de una tarea el tiempo disponible puede provocar un sesgo en favor de uno u otro mecanismo. O incluso es posible que el individuo estratégicamente cambie los criterios en favor o en perjuicio de uno de los mecanismos.

El modelo de Malmberg y Xu (2007) y Xu y Malmberg (2007; véase también Malmberg, 2008) nos dice que los procesos de familiaridad proporcionan información más rápidamente que los procesos de recuperación. Estos autores siguen en su modelo las ideas de Atkinson y Juola (1974; Atkinson, Herrmann y Westcourt, 1974). La persona tiene en cuenta en primer lugar la familiaridad de la clave del par examinado comparándola con un criterio subjetivo. Si esta familiaridad no sobrepasa el criterio, la respuesta es negativa (nuevo). Si sobrepasa el criterio se inician procesos de recuperación. Si estos recuperan detalles episódicos, son comparados con el ítem; si son similares al mismo, la respuesta es positiva (antiguo) y, si no son similares, la respuesta es negativa (nuevo). Si los procesos de recuperación fallan, las personas tienden a dar suposiciones positivas.

Visto todo lo anterior, ¿cómo podríamos explicar que la precisión de las estimaciones de C4 no sea superior a C1 ante el reconocimiento en Weaver y Kelemen (2003)? Pues bien, al contrario de lo que suponen estos autores, C1 y C4 son diferentes en contexto, pero no está garantizado que lo sean en *procesamiento*.

La existencia de los dos procesos en la tarea de reconocimiento hace que, para que el juicio C4 sea predictivo, el individuo en cada ítem tiene que hacer el mismo uso de una y otra información (familiaridad y recuperación asociativa) tanto en el juicio como en la prueba. Pensamos que en los juicios de la condición C4 la persona se basa principalmente en los procesos de recuperación (el heurístico de la recuperabilidad que vimos arriba). Por este motivo, el procesamiento utilizado en C4 es similar al de C1. En otras palabras, aunque no haya equivalencia o correspondencia aparente entre los contextos de C4 y C1, es posible que sí la haya en términos de procesamientos implicados¹⁰.

7.- OBJETIVOS DE INVESTIGACIÓN.

El objetivo de esta tesis es realizar una nueva evaluación empírica de la hipótesis de la *Supervisión Apropiaada para la Transferencia* (SAT) como explicación teórica de la precisión de los "juicios de aprendizaje".

En los Experimentos 1 y 2 vamos a estudiar el efecto sobre la precisión de las estimaciones de las personas de dos tipos de claves para los juicios de aprendizaje: estímulo solo (ES) (C1 en Weaver y Kelemen, 2003) y estímulo más 4 alternativas de respuesta (ER) (C4 en Weaver y Kelemen, 2003), y dos tipos de pruebas de memoria: recuerdo (RD) y reconocimiento (RN). Así, cuando presentamos el estímulo solo para los juicios, utilizamos el mismo contexto que para la prueba de recuerdo. Por otro lado, cuando presentamos el estímulo más 4 alternativas de respuesta para los juicios, utilizamos el mismo contexto que para la prueba de reconocimiento.

¹⁰ La distinción entre procesamiento y contexto en la *Supervisión Apropiaada para la Transferencia* (SAT) ha sido también sugerida por Roediger y Guynn (1996).

Como hemos visto antes, en C4 es posible que los participantes se basen primordialmente en el componente de recuperación tanto en los juicios como en la prueba de reconocimiento asociativo. Este hecho puede ser debido a que todas las alternativas de respuesta en C4 y reconocimiento estaban formadas por otros objetivos de la prueba. Como hemos visto, al estar igualadas en familiaridad pueden inducir a los participantes a emplear primordialmente la recuperabilidad como propiedad discriminativa. Para evitar esto, en nuestros Experimentos 1 y 2 intentamos potenciar el papel de la familiaridad introduciendo alternativas externas: el 50% de las alternativas serán otros objetivos de la prueba y el 50 % restante serán palabras externas al experimento (en juicios y pruebas). Así, la persona puede basar sus juicios y respuestas de reconocimiento tanto en el componente de recuperación como en el de familiaridad.

Por otro lado, debemos añadir que cuando aumenta el número de alternativas en la prueba de reconocimiento asociativo, la prueba de reconocimiento se asemeja cada vez más a una prueba de recuerdo (Davis, Sutherland y Judd, 1961). En consecuencia, según aumentamos las alternativas, el contexto de reconocimiento se asemeja más al contexto de recuerdo (Dunlosky y Nelson, 1997). Para que haya una mayor diferenciación entre las pruebas de recuerdo y reconocimiento, en nuestros Experimentos 1 y 2 reducimos el número de alternativas de 7 a 4 con respecto al experimento de Weaver y Kelemen (2003), tanto en juicios como en reconocimiento.

Si la hipótesis SAT es correcta, juicios y pruebas de memoria deben interactuar. En la prueba de recuerdo deberíamos esperar una mayor precisión de los juicios realizados ante el estímulo solo que ante el estímulo más las respuestas. Y en la prueba de reconocimiento deberíamos esperar una mayor precisión de los juicios realizados ante el par estímulo+respuesta que ante el estímulo solo. El objetivo

principal de los Experimentos 1 y 2 es evaluar empíricamente si se produce esta interacción.

Como veremos más adelante, el objetivo del Experimento 3 es intervenir en lo posible sobre los procesos de recuperación y familiaridad en la prueba de reconocimiento.

8.- MEDIDA DE LA PRECISIÓN DE LOS JUICIOS DE APRENDIZAJE.

Antes de pasar a describir los experimentos, conviene que hagamos algunas precisiones sobre nuestras medias de metamemoria. Las estimaciones de las personas son evaluadas mediante medidas de precisión absoluta y relativa (Nelson y Dunlosky, 1991 y Dunlosky y Nelson, 1992).

La precisión relativa de los juicios de aprendizaje también se denomina *resolución* (p. ej., Van Overschelde y Nelson, 2006). La resolución es el grado en que una persona puede predecir la diferencia de probabilidad de recuperación de un ítem con respecto a otro ítem de la misma lista estudiada. Según Nelson (1984), el mejor índice para medir la resolución es la correlación gamma de Goodman-Kruskal (G) entre los juicios de aprendizaje y la recuperación posterior. Como suele ser el caso con los índices de correlación, su rango va de -1.0 a +1.0, donde 0 representa una precisión nula y +1.0 representa una precisión perfecta de las predicciones.

Aunque existe en la actualidad un debate intenso sobre cómo medir la resolución metamnemónica (véase, por ejemplo, Barrett, Dienes y Seth, 2013; Benjamin y Díaz, 2008; Maniscalco y Lau, 2012; Masson y Rotello, 2009), con el fin de poder comparar nuestros resultados con los obtenidos en la investigación que nos

sirve de contraste, usaremos a lo largo de esta tesis la Gamma de Goodman-Kruskal, que fue la medida utilizada por Weaver y Kelemen (2003). No obstante, conviene decir aquí que algunos análisis exploratorios con otras medidas -no recogidos en esta tesis- no han arrojado diferencias reseñables (Ruiz y Arroyo, 2015).

Por su parte, en cuanto a la precisión absoluta o *calibración* (Van Overschelde y Nelson, 2006), nos informa del grado en que una persona se aproxima en su predicción a su nivel global o promedio de rendimiento en una prueba posterior. Es decir, la correspondencia absoluta entre los juicios emitidos para un grupo de ítems y sus posibles respuestas correctas, por ejemplo: si un grupo de ítems recibe un promedio de estimación de acierto del 40%, la persona obtendrá una calibración absoluta o ajuste perfecto si su rendimiento en la prueba posterior es del 40% de aciertos. Obsérvese que esto equivale a decir que la predicción específica de un ítem concreto se ajusta a su particular probabilidad de recuperación. En el presente trabajo la precisión absoluta se denominará *ajuste*.

Uno de los principales problemas para medir el ajuste entre los juicios y el rendimiento es la escala de medida, debido a que los juicios se solicitan en una escala ordinal o dicotómica (Koriat, 1997). Esto impide utilizar un índice de correlación adecuado como, por ejemplo, la correlación de Pearson. Nelson y Dunlosky (1991) y Dunlosky y Nelson (1992) muestran la precisión absoluta o ajuste mediante curvas de calibración. Se trata de una representación gráfica donde en el eje de abscisas aparece la magnitud de los juicios de aprendizaje agrupada en intervalos (0-10%, 11-20%...) y en el eje de ordenadas aparece el porcentaje de aciertos en la prueba correspondiente a cada intervalo de juicios. En la presentación de nuestros datos de calibración absoluta seguiremos en líneas generales su procedimiento de cálculo.

9.- SECCIÓN EXPERIMENTAL.

9.1- EXPERIMENTO 1.

En este experimento vamos a utilizar el mismo contexto para realizar los juicios y las pruebas. Así, el contexto para emitir los juicios de aprendizaje cuando presentamos el estímulo solo (ES); es el mismo contexto de la prueba de recuerdo. Mientras que; el contexto para emitir los juicios cuando presentamos el estímulo más las alternativas de respuesta (ER), es el mismo de la prueba de reconocimiento. Por este motivo, debemos esperar que el procesamiento requerido para los juicios ES, sea más similar al procesamiento requerido en la prueba de recuerdo. Y, por otro lado, también debemos esperar que el procesamiento requerido para los juicios ER, sea más similar al procesamiento requerido en la prueba de reconocimiento. Por tanto, en línea con los supuestos básicos de la hipótesis de la *Supervisión Apropriada para la Transferencia* (SAT), se debe producir interacción entre las variables *tipo de juicio* y *tipo de prueba* en la precisión de las estimaciones: cuando los procesamientos en una y otra tarea se asemejen, la precisión de las predicciones será mayor. El objetivo del Experimento 1 es comprobar si se produce esta interacción. Para ello, se van a cruzar las dos variables en un diseño completamente intrasujeto.

MÉTODO

PARTICIPANTES

Participaron 63 estudiantes universitarios de la Facultad de Psicología de la UNED (49 mujeres y 14 hombres), con una edad comprendida entre los 22 y 59 años (M= 38, DT= 10), para obtener créditos en la asignatura *Psicología de la Memoria*.

MATERIALES Y APARATOS

Los ítems eran 120 pares de palabras no relacionadas formados por sustantivos seleccionados de la lista de la Universidad de Valencia de Algarabel, Ruiz, y San Martín (1988). Los ítems se seleccionaron teniendo en cuenta la variable frecuencia de las palabras, 40 pares estaban formados por palabras de alta frecuencia ($N_{\text{por}_500K} \geq 160$), 40 pares estaban formados por palabras de media frecuencia ($20 \leq N_{\text{por}_500K} \leq 50$) y los 40 pares restantes por palabras de baja frecuencia ($N_{\text{por}_500K} \leq 7$). Los 120 pares de palabras se dividían en 4 bloques de 30 ítems. Cada bloque era mostrado de forma grupal a los participantes mediante presentaciones de Microsoft PowerPoint con proyector.

Las palabras que se utilizaban como distractores en la prueba de reconocimiento eran diferentes a las presentadas durante los juicios, tomando 120 palabras adicionales como distractores para los juicios y otras 120 palabras adicionales para la prueba de reconocimiento.

PROCEDIMIENTO Y DISEÑO

El diseño experimental comprende las siguientes variables: *tipo de juicio* (estímulo solo (ES) vs. estímulo más respuestas (ER)) y *tipo de prueba de memoria* (recuerdo (RD) vs. reconocimiento (RN)).

Los juicios de aprendizaje se emitían en dos condiciones: presentando la palabra estímulo sola (ES) o presentando la palabra estímulo junto con 4 alternativas (ER). Dicha condición era una manipulación intra-sujeto. Las pruebas de memoria eran también una manipulación intra-sujeto: recuerdo (RD), donde se le muestra al participante la palabra estímulo sola, debiendo recordar la palabra objetivo, y reconocimiento (RN), donde se presenta la palabra estímulo junto con 4 alternativas,

debiendo seleccionar una de las mismas. Tanto en los Juicios ER como en la prueba de reconocimiento las 4 alternativas eran la respuesta correcta, una palabra estudiada (un término respuesta de otro par elegido al azar) y dos palabras externas.

Las variables dependientes eran: la *Magnitud de los Juicios de Aprendizaje*, el *Rendimiento en las Pruebas* de memoria y la *Precisión de los Juicios de Aprendizaje*.

Los ítems son asignados de forma aleatoria a las dos condiciones de juicios, 60 ítems para la presentación de la palabra estímulo sola y los 60 ítems restantes para la presentación de la palabra estímulo junto con 4 alternativas (palabra objetivo, palabra distractora seleccionada al azar entre los objetivos y dos palabras externas no presentadas anteriormente). Los 60 ítems de cada condición de juicios se asignan de forma aleatoria a las dos condiciones de prueba de memoria, el 50% para la prueba de recuerdo y el 50% restante para la prueba de reconocimiento. Los ítems son asignados de forma aleatoria a los cuatro bloques, el primer bloque estaba formado por 8 ítems ES-RD, 7 ítems ES-RN, 7 ítems ER-RD y 8 ítems ER-RN; el segundo bloque estaba formado por 7 ítems ES-RD, 8 ítems ES-RN, 8 ítems ER-RD y 7 ítems ER-RN; el tercer bloque se confeccionó de la misma forma que el primero y el cuarto bloque de la misma forma que el segundo.

Las tareas de estudio y juicio se intercalaron por bloque. Comenzando por el primer bloque, los participantes realizaban una *Tarea de Estudio* de pares asociados. Los ítems eran presentados en la pantalla durante 6 seg. (p. ej., guerra-NÚMERO). La palabra estímulo (guerra) aparecía en minúsculas en la parte superior izquierda y la palabra objetivo (NÚMERO) aparecía en mayúsculas en el centro de la pantalla. Finalizada la tarea de estudio se presentaba una tarea distractora (suma de cifras de varios dígitos con lápiz y papel) de 2 minutos de duración. Posteriormente, se

realizaban los *Juicios de Aprendizaje* de este primer bloque. En una nueva presentación grupal con proyector, en la pantalla aparecía durante 8 segundos la palabra estímulo sola o la palabra estímulo más 4 alternativas, seguida de la siguiente pregunta: *¿Hasta qué punto está seguro/a de que cuando se le presente la primera palabra del par dentro de unos diez minutos será capaz de acertar su segunda palabra?*. En la parte superior derecha de la pantalla aparecía el número del ítem; para responder, se proporcionaba a los participantes una hoja de respuestas con la relación de ítems, donde debían puntuar a cada uno su grado de seguridad de acierto, escribiendo un número en una escala de 0 a 100 (p. ej., 87% de seguridad). Los ítems se presentaban con la siguiente restricción de secuencia para la emisión de los Juicios de Aprendizaje: los primeros 10 pares presentados durante el estudio eran aleatorizados y presentados para los JA, y así los siguientes 10 ítems y los 10 restantes del bloque.

Una vez finalizado el estudio y emisión de JA del primer bloque, se repetía el procedimiento para los bloques segundo, tercero y cuarto. Y luego, al finalizar los cuatro bloques de estudio-juicio, los participantes realizaban una tarea distractora (multiplicar y dividir) durante 5 minutos.

Finalmente, se realizaban las *Pruebas de Memoria*, presentando de forma aleatoria los 10 primeros ítems de la fase de estudio, luego los 10 segundos y 10 terceros del primer bloque, y 10 primeros, 10 segundos y 10 terceros del segundo bloque, y así hasta el cuarto bloque. Para las pruebas los ítems eran presentados mediante el proyector. Para la prueba de recuerdo se presentaba el estímulo solo durante 8 segundos, debiendo escribir el participante la palabra objetivo en la hoja de respuestas. Para la prueba de reconocimiento aparece la palabra estímulo junto con 4 alternativas durante 8 segundos, debiendo marcar en la hoja de respuestas la

alternativa A, B, C ó D. Es importante destacar que desde el comienzo de la prueba se comunica a los participantes que deben responder a todos los ítems, aclarándoles que habrá cierta permisividad (sinónimos o palabras similares puntuarán en recuerdo). Los ensayos de uno y otro tipo de prueba se sucedían al azar sin solución de continuidad.

Las pruebas se realizaron de forma grupal en el Salón de Actos de la Facultad de Psicología de la UNED, en 3 sesiones diferentes para permitir una mayor separación espacial entre los participantes, con un número similar de participantes en cada una. Para cada sesión los ítems se presentaban en una aleatorización diferente con las restricciones que se han comentado arriba.

RESULTADOS

RENDIMIENTO EN LAS PRUEBAS DE MEMORIA

El rendimiento de memoria se calculó mediante el porcentaje de respuestas acertadas por cada participante en las pruebas de recuerdo y reconocimiento. En la corrección de la prueba de reconocimiento se introduce la fórmula de corrección por adivinación (Macmillan y Creelman, 1991):

$$R = A - (E/k-1)$$

(R= Rendimiento prueba, A= Aciertos, E= Errores, K= n° de alternativas)

El promedio del rendimiento en las pruebas de memoria para cada condición experimental puede verse en la Tabla 1.1.

Se realizó un ANOVA 2 (*tipo de juicio*) X 2 (*tipo de prueba*)¹¹ de medidas repetidas del promedio del rendimiento en las pruebas. En un primer análisis se incluyeron las variables *sesión* y *bloque*, pero estas variables procedimentales no presentaron efectos ni interacciones sistemáticas, por lo que no serán incluidas en la descripción de los resultados. No obstante, pueden verse los resultados completos en el Apéndice 1.1.

Se produce un efecto significativo del *tipo de juicio*, $F(1, 60) = 21.05$, $MCE = 366.58$, $p < .05$, $\eta^2 = 0.26$. El rendimiento en las pruebas es superior cuando presentamos el estímulo junto con varias alternativas de respuesta ($M=31$, $DT=35$) que cuando presentamos el estímulo solo ($M=26$, $DT=31$). Esto sugiere que la presentación para los juicios del par estudiado con los distractores (ER), actuó como una sesión de estudio adicional de los ítems. También se produce un efecto significativo del *tipo de prueba*, $F(1, 60) = 226.16$, $MCE = 1025.89$, $p < .05$, $\eta^2 = 0.79$. En la prueba de reconocimiento ($M=44$, $DT=32$) el rendimiento es más elevado que en la prueba de recuerdo ($M=13$, $DT=33$). Hay que destacar, que se produce un bajo rendimiento en la prueba de recuerdo: el 30% de los participantes no recuerda ninguna palabra en la condición ES y el 24% de los participantes no recuerda ninguna palabra en la condición ER. Por último, la interacción entre las variables *tipo de juicio* y *tipo de prueba* estuvo próxima a la significación, $F(1, 60) = 3.40$, $MCE = 463.73$, $p < .10$, $\eta^2 = 0.05$, ya que el efecto del *tipo de juicio* fue algo más grande en recuerdo que en reconocimiento.

¹¹ Siempre que en el análisis de varianza se incluyan factores con medidas repetidas, se utilizará la corrección de Greenhouse-Geisser (Weinfurt, 2000; Howell, 2013).

	RENDIMIENTO		JUICIOS	
	RD	RN	RD	RN
ES	9 ₍₂₉₎	43 ₍₂₇₎	30 ₍₃₂₎	29 ₍₃₂₎
ER	17 ₍₃₈₎	45 ₍₂₇₎	51 ₍₃₈₎	51 ₍₃₈₎

Tabla 1.1. Promedio y desviación típica de la magnitud de los juicios y del porcentaje de aciertos en las pruebas de memoria (ES: estímulo solo, ER: estímulo más respuestas, RD: recuerdo, RN: reconocimiento).

MAGNITUD DE LOS JUICIOS DE APRENDIZAJE

El promedio de la magnitud de los juicios de aprendizaje para cada condición experimental puede verse en la Tabla 1.1.

Se realizó un ANOVA 2 (*tipo de juicio*) X 2 (*tipo de prueba*) de medidas repetidas del promedio de la magnitud de los juicios de aprendizaje. Una vez más, no se incluyen en la descripción de los resultados las variables *sesión* y *bloque*, ya que no presentaron efectos ni interacciones sistemáticas, los resultados completos se pueden consultar en el Apéndice 1.2.

Los resultados mostraron un efecto significativo del *tipo de juicio*, $F(1, 60) = 136.70$, $MCE = 886.57$, $p < .05$, $\eta^2 = 0.69$. La magnitud de los juicios de aprendizaje es superior en la condición ER ($M=51$, $DT=38$) que en la condición ES ($M=29$, $DT=32$). Como era de esperar, el tipo de prueba que recibirán los participantes no afecta a la magnitud de los juicios, $F < 1$, ya que la clasificación de los ítems por el *tipo de prueba* es totalmente desconocida para los participantes en el momento del juicio. Tampoco se produce interacción entre las variables *tipo de juicio* y *tipo de prueba*, $F < 1$.

DISTRIBUCIÓN DE LOS JUICIOS DE APRENDIZAJE

Para analizar de una forma más detallada cómo se distribuyen los juicios de aprendizaje (es decir, cómo usan las personas la escala de valoración cuando realizan sus estimaciones), siguiendo el procedimiento utilizado por Dunlosky y Nelson (1994), examinaremos el porcentaje de ítems que recibe cada nivel de valoración de los juicios.

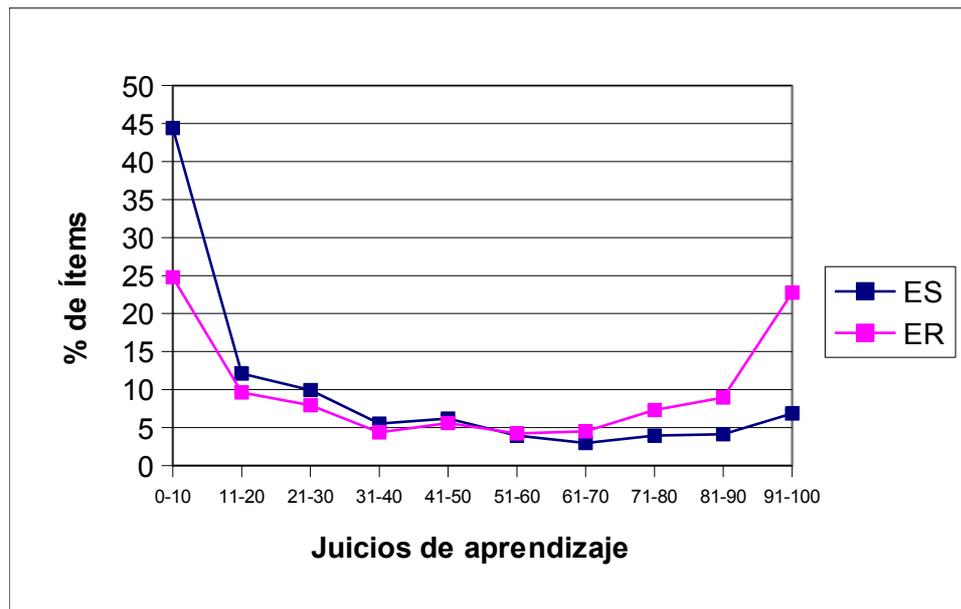


Gráfico 1.1. Distribución de los juicios de aprendizaje. Porcentaje de respuestas para cada nivel de juicio (ES: estímulo solo. ER: estímulo más respuestas).

Como podemos observar en el Gráfico 1.1, la distribución de los juicios cambia en función del estímulo presentado para su emisión. En los juicios ES la distribución es asimétrica positiva¹²; mientras que en los juicios ER se da una distribución bimodal. Es bien sabido que esta última forma de distribución es característica de los juicios demorados (Dunlosky y Nelson, 1994), como es nuestro caso. La diferencia entre ambas distribuciones es estadísticamente significativa: $\chi^2(9) = 17.92$ $p < .05$.

¹² Una distribución es asimétrica positiva si la "cola" a la derecha de la media es más larga que la de la izquierda, es decir, si hay valores más separados de la media hacia la derecha.

PRECISIÓN DE LOS JUICIOS DE APRENDIZAJE: RESOLUCIÓN

Para cada participante se calcularon 4 correlaciones gamma entre la magnitud de los juicios de aprendizaje y el rendimiento en las pruebas, una por condición experimental. Los promedios de dichas correlaciones pueden verse en la Tabla 1.2.

Se realizó un ANOVA 2 (*tipo de juicio*) X 2 (*tipo de prueba*) de medidas repetidas de la correlación gamma. Se excluye del análisis la variable *bloque* al corresponder un número demasiado reducido de ítems. También aquí la variable *sesión* no presentó efectos significativos sistemáticos, por lo que no se incluye en la descripción de los resultados (Ver Apéndice 1.3).

Se produce un efecto significativo del *tipo de prueba*, $F(1, 34) = 86.39$, $MCE = 0.06$, $p < .05$, $\eta^2 = 0.72$; las estimaciones son más acertadas en la prueba de recuerdo ($M=0.80$, $DT=0.25$) que en la prueba de reconocimiento ($M=0.45$, $DT=0.31$). Este efecto replica los resultados obtenidos por Begg y otros. (1989), Thiede y Dunlosky (1994) y Weaver y Kelemen (2003). El efecto del tipo de juicio no es significativo, $F(1, 34) = 2.25$, $MCE = 0.07$, $p < .10$, $\eta^2 = 0.06$.

	RD	RN
ES	0.81 (0.25)	0.37 (0.29)
ER	0.78 (0.25)	0.53 (0.34)

Tabla 1.2. Promedio de la correlación gamma y desviación típica (ES: estímulo solo, ER: estímulo más respuestas, RD: recuerdo, RN: reconocimiento).

Se produce interacción significativa entre las variables *tipo de juicio* y *tipo de prueba*, $F(1, 34) = 4.39$, $MCE = 0.04$, $p < .05$, $\eta^2 = 0.11$. Como podemos observar en la Tabla 1.2, en la prueba de recuerdo, la precisión de las estimaciones es similar en

ambas condiciones de juicios; sin embargo, en la prueba de reconocimiento, la precisión de las estimaciones es superior cuando presentamos ER que cuando presentamos ES. Un análisis específico en la prueba de recuerdo con *t* de Student no muestra diferencias significativas ES-ER, $t(38)=0.56$, $p>.10$. Por el contrario, un análisis específico en la prueba de reconocimiento con *t* de Student muestra diferencias significativas ES-ER, $t(58)^{13}=-3.07$, $p<.05$.

PRECISIÓN DE LOS JUICIOS DE APRENDIZAJE: AJUSTE

El grado de ajuste entre las predicciones de los participantes y su rendimiento en las pruebas de memoria se muestra mediante curvas de calibración.

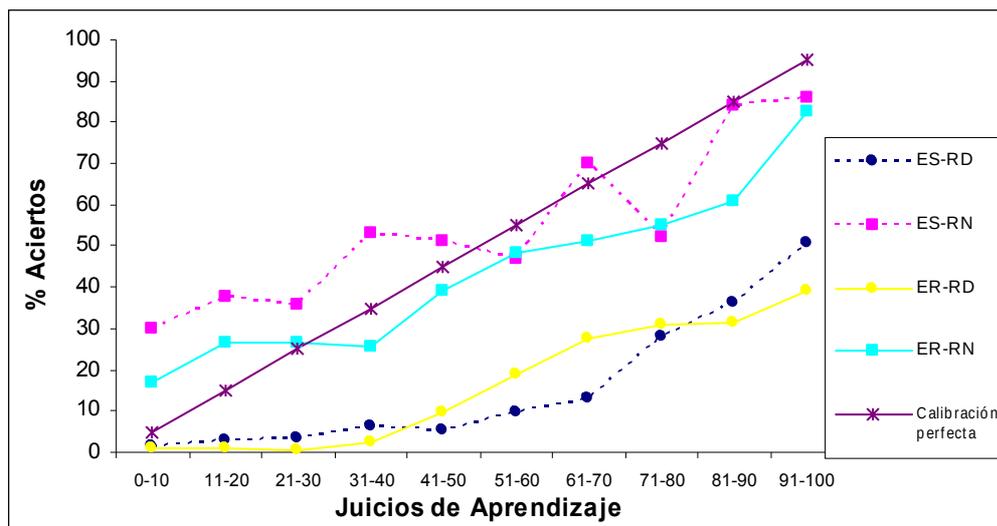


Gráfico 1.2. Curvas de calibración para cada condición experimental (ES: estímulo solo, ER: estímulo más respuestas, RD: recuerdo. RN: reconocimiento).

Para su cálculo se sigue aquí el mismo procedimiento de Nelson y Dunlosky (1991) y Dunlosky y Nelson (1992): en cada condición experimental la magnitud de los juicios de aprendizaje se agrupó en intervalos, anotando en cada intervalo el porcentaje correspondiente de aciertos en la prueba de memoria asociada. El

¹³ En la prueba de reconocimiento hay más grados de libertad porque en la prueba de recuerdo se producen más gammas indeterminadas. Una gamma es indeterminada cuando no es posible su cálculo, bien porque se emite siempre el mismo juicio o porque se aciertan o fallan todos los ítems. En este caso se producen las siguientes gammas indeterminadas: ESRD=19, ESRN=4, ERRD=15 y ERRN=0.

rendimiento de la prueba de reconocimiento se obtiene mediante la fórmula de corrección por adivinación expresada anteriormente. Los resultados se muestran en el Gráfico 1.2, en el que la diagonal muestra la calibración perfecta.

Para las cuatro curvas de calibración de las condiciones experimentales se han calculado sus correspondientes rectas de regresión, que se muestran en la Tabla 1.3.

	RD	RN
ES	$y=-9.33+0.50x$	$y=25.63+0.58x$
ER	$y=-7.72+0.48x$	$y=11.24+0.64x$

Tabla 1.3. Rectas de regresión lineal. Calibración perfecta:
 $y=0+1x$

La regresión lineal es estadísticamente significativa en las cuatro rectas (Ver Apéndice 1.4), siendo la correlación muy alta en todas ellas. El nivel de ajuste es más elevado en los juicios ER que en los juicios ES, como podemos observar en la Tabla 1.4.

	RD	RN
ES	0.81	0.81
ER	0.93	0.93

Tabla 1.4. Coeficiente de determinación R^2 de las rectas de regresión.

Un análisis de las rectas de regresión, comparándolas de dos en dos, mediante el estadístico ZPF¹⁴ propuesto por Raghunathan, Rosenthal y Rubin (1996; también Ruiz y Arroyo, 2014; Steiger, 1980) no mostró diferencias significativas en los ajustes (Ver Apéndice 1.4). Aunque las diferencias se aproximan a la significación cuando comparamos las pruebas en función de los juicios: ESRD-ERRD, $ZPF(N=10)=1.40$, $p<0.08$ y ESRN-ERRN, $ZPF(N=10)=1.34$, $p<0.09$.

¹⁴ Cuando necesitamos realizar una prueba de significación estadística de la diferencia entre dos correlaciones relacionadas no solapadas, y otros (1996) recomiendan calcular la puntuación ZPF.

DISCUSIÓN

Es bien sabido que, salvo circunstancias especiales (p. ej., Thomson y Tulving, 1970; Tulving y Thomson, 1971), el rendimiento en las pruebas de memoria suele ser superior en reconocimiento que en recuerdo. Esto es conocido por la población y forma parte de la teoría sobre metamemoria compartida en nuestra sociedad, por lo que actuaría como clave normativa (o clave intrínseca en la terminología de Koriat, 1997) para emitir juicios más altos ante ER. Además, una mayor dificultad del juicio mismo ante ES actuaría como clave extrínseca y/o mnemónica para emitir juicios más bajos. Por todo ello, debemos esperar que la magnitud de los juicios ER sea superior a los juicios ES. Y efectivamente, esto ocurre así. Además, vemos cómo las estimaciones de las personas están en consonancia con el rendimiento final en las pruebas. De acuerdo con esto, la distribución de los juicios cambia en función del estímulo presentado como contexto de juicio, lo cual sugiere que las personas están respondiendo de forma diferente ante un contexto y otro. En resumen, tanto los datos sobre la magnitud de los juicios de aprendizaje como su distribución invitan a pensar que las personas utilizan procesamientos diferentes de la información dependiendo de los estímulos presentados como contexto de juicio.

Ahora bien, el resultado más importante de este experimento es que la precisión relativa de los juicios de aprendizaje, medida por la correlación gamma, muestra una interacción entre las variables *tipo de juicio* y *tipo de prueba*. Estos resultados están de acuerdo con las predicciones de la SAT y difieren claramente de los informados por Weaver y Kelemen (2003). En la prueba de reconocimiento la resolución es superior ante ER que ante ES, lo cual sugiere que, cuando presentamos el estímulo junto con cuatro alternativas de respuesta (ER), el procesamiento de la

información provocado durante los juicios tiene más similitud con el procesamiento requerido en la prueba de reconocimiento que el provocado por ES.

También hay que destacar que, según la hipótesis SAT, para la prueba de recuerdo debemos esperar que la precisión de las estimaciones sea inferior cuando presentamos el estímulo más las respuestas (ER) en los juicios que cuando presentamos el estímulo solo (ES) (p. ej., Nelson y Dunlosky, 1991; Begg y otros, 1989). Sin embargo, este hecho no se produce. Como podemos observar en la Tabla 1.1, esto podría deberse a la baja tasa de aciertos en la prueba de recuerdo, que conlleva a su vez una escasa varianza en la prueba. Cabe sospechar que en la prueba de recuerdo nos hemos acercado a un efecto de suelo.

9.2- EXPERIMENTO 2.

En consonancia con las predicciones de la hipótesis SAT, en el experimento anterior se produce interacción entre las variables *tipo de juicio* y *tipo de prueba*. Sin embargo, el bajo rendimiento en recuerdo puede haber producido una escasa varianza en esta prueba, quizá debido a un efecto de suelo. Por este motivo y ante la importancia del resultado obtenido en el Experimento 1 respecto al efecto del contexto de juicio de aprendizaje sobre su resolución como predicción de reconocimiento, el objetivo principal que nos planteamos en el Experimento 2 es replicar los resultados anteriores con un rendimiento superior en recuerdo.

Con esta finalidad, se introducen algunas modificaciones con respecto al Experimento 1. En primer lugar, se reduce el número de ítems presentados así como el tiempo de presentación de los mismos en la fase de estudio, que pasa de 6 seg. a 5 seg.. Con esto conseguimos reducir el tiempo de duración de la fase de prueba y la carga de memoria o la interferencia inter- e intralista, exigiendo un menor esfuerzo a los participantes (Lohnas, Polyn y Kahana, 2015). En segundo lugar, se eliminan los ítems de media frecuencia, para aumentar la varianza de rendimiento intrasujeto (p. ej., Criss y Shiffrin, 2004; Malmberg, Holden y Shiffrin, 2004). En tercer lugar, durante la fase de estudio, los ítems que reciben la prueba de recuerdo son presentados 2 veces. Por último, se aumenta el tiempo de respuesta en la prueba de recuerdo que pasa de 8 seg. a 11 seg.

La fiabilidad en el Experimento 1 es muy alta. Para los juicios de aprendizaje se obtiene un índice de fiabilidad¹⁵ R_{xx} de 0.93 y para las pruebas de memoria un índice de fiabilidad R_{xx} de 0.96. Si reducimos el número de ítems de la prueba de 120 a 80 ítems, la fiabilidad estimada¹⁶ sería de $R_{xx}=0.90$ y $R_{xx}=0.95$, respectivamente.

¹⁵ Para calcular la fiabilidad se ha utilizado el procedimiento de las dos mitades de Spearman-Brown.

¹⁶ Para calcular la fiabilidad estimada se ha utilizado la ecuación de Spearman-Brown que pone en

Seguimos manteniendo, por tanto, un elevado índice de fiabilidad en ambos casos.

MÉTODO

PARTICIPANTES

Participaron 53 estudiantes universitarios de la Facultad de Psicología de la UNED (40 mujeres y 13 hombres), con una edad comprendida entre los 21 y 62 años ($M= 37.11$, $DT= 9.81$), para obtener créditos en la asignatura *Psicología de la Memoria*.

MATERIALES Y APARATOS

Los ítems eran 80 pares de palabras no relacionadas formados por sustantivos seleccionados de la lista de la Universidad de Valencia de Algarabel y otros (1988). Empleando las palabras de alta y baja frecuencia del Experimento 1. Los 80 pares de palabras se dividían en 2 bloques de 40 ítems. La forma de presentación de los ítems a los participantes era la misma que en el Experimento 1.

Como en el experimento anterior, las palabras que se utilizaron como distractores en la prueba de reconocimiento eran diferentes a las presentadas durante los juicios, tomando 80 palabras adicionales como distractores para los juicios ER y otras 80 palabras adicionales para la prueba de reconocimiento.

PROCEDIMIENTO Y DISEÑO

El procedimiento y diseño experimental es el mismo que en el Experimento 1, salvo algunos cambios que se exponen a continuación.

Se asignaron los ítems de forma aleatoria a las dos condiciones de juicios, 40

relación la fiabilidad y el número de elementos (Ver Apéndice 2.5).

ítems para la condición de juicios ES y los 40 ítems restantes para la condición ER. Los 40 ítems de cada condición de juicios se asignan de forma aleatoria a las dos condiciones de prueba de memoria, el 50% para la prueba de recuerdo y el 50% restante para la prueba de reconocimiento. Los ítems se asignaron de forma aleatoria a los dos bloques. Cada bloque estaba formado por 10 ítems ES-RD, 10 ítems ES-RN, 10 ítems ER-RD y 10 ítems ER-RN.

Comenzando por el primer bloque, los participantes realizaban una *Tarea de Estudio* de pares asociados, los ítems eran presentados en la pantalla durante 5 seg. de la misma forma que en el Experimento 1. Los ítems que iban a recibir la prueba de recuerdo eran presentados dos veces durante la fase de estudio, con la única restricción de que antes de cada repetición debía haber al menos 5 ítems no repetidos. Por tanto, incluidas las repeticiones cada bloque de estudio contenía 60 presentaciones de pares de palabras. Finalizada la tarea de estudio se presentaba una tarea distractora (suma de cifras de varios dígitos con lápiz y papel) de 2 minutos de duración. Posteriormente, se realizaban los *Juicios de Aprendizaje* de este primer bloque con el mismo procedimiento que en el Experimento 1. Una vez finalizado el estudio y emisión de JA del primer bloque, se repetía el procedimiento con el segundo bloque, incluida la tarea distractora.

Finalmente, se realizaban las *Pruebas de Memoria*. Con el mismo procedimiento que en el Experimento 1, pero con algunas modificaciones. Para la prueba de recuerdo se presentaba el estímulo solo durante 11 segundos, debiendo escribir el participante la palabra objetivo en la hoja de respuestas. Para la prueba de reconocimiento aparece la palabra estímulo junto con las 4 alternativas de respuesta durante 8 segundos, debiendo escribir la palabra correspondiente al objetivo. Los ensayos de uno y otro tipo se suceden al azar.

Las pruebas se realizaron de forma grupal en el salón de actos de la Facultad de Psicología de la UNED, en 2 sesiones diferentes, con un número similar de participantes en cada una. En la segunda sesión los ítems se aleatorizaron de nuevo, formando las parejas de palabras de forma totalmente diferente a la primera sesión.

RESULTADOS

RENDIMIENTO EN LAS PRUEBAS DE MEMORIA

El rendimiento de memoria se calculó de la misma forma que en el Experimento 1: en reconocimiento se utiliza también la fórmula de corrección por adivinación. El promedio del rendimiento en las pruebas para cada condición experimental puede verse en la Tabla 2.1. Vemos que se produce un aumento del rendimiento de la prueba de recuerdo, que pasa de un promedio del 13% al 31% de aciertos con respecto al experimento anterior. Por tanto, se logra uno de los objetivos de este experimento.

Se realiza el mismo análisis que en el Experimento 1. Pueden verse los resultados completos en el Apéndice 2.1.

Replicando los hallazgos del Experimento 1, se produce un efecto significativo del *tipo de juicio*, $F(1, 51) = 32.43$, $MCE = 270.06$, $p < .05$, $\eta_p^2 = 0.39$. El rendimiento en las pruebas es superior cuando presentamos ER ($M=44$, $DT=37$) que cuando presentamos ES ($M=35$, $DT=34$). También se replica el efecto del *tipo de prueba*, $F(1, 51) = 113.79$, $MCE = 315.23$, $p < .05$, $\eta_p^2 = 0.69$. En la prueba de reconocimiento ($M=49$, $DT=25$) el rendimiento es más elevado que en la prueba de recuerdo ($M=31$, $DT=46$). Por último, la interacción entre las variables *tipo de juicio* y *tipo de prueba* es significativa, $F(1, 51) = 6.09$, $MCE = 178.93$, $p < .05$, $\eta_p^2 = 0.11$,

ya que el efecto del *tipo de juicio* fue mayor en reconocimiento que en recuerdo.

	RENDIMIENTO		JUICIOS	
	RD	RN	RD	RN
ES	28 ₍₄₅₎	43 ₍₂₄₎	46 ₍₃₉₎	30 ₍₃₂₎
ER	33 ₍₄₇₎	55 ₍₂₇₎	68 ₍₃₅₎	53 ₍₃₉₎

Tabla 2.1. Promedio y desviación típica de la magnitud de los juicios y del porcentaje de aciertos en las pruebas de memoria (ES: estímulo solo, ER: estímulo más respuestas, RD: recuerdo, RN: reconocimiento).

MAGNITUD DE LOS JUICIOS DE APRENDIZAJE

El promedio de la magnitud de los juicios de aprendizaje para cada condición experimental puede verse en la Tabla 2.1. Se realizó el mismo análisis que en el Experimento 1. Pueden verse los resultados completos en el Apéndice 2.2.

Replicando los resultados del Experimento 1, hay un efecto significativo del *tipo de juicio*, $F(1, 51) = 120.50$, $MCE = 431.53$, $p < .05$, $\eta_p^2 = 0.70$. La magnitud de los juicios de aprendizaje es superior en la condición ER ($M=60$, $DT=37$) que en la condición ES ($M=38$, $DT=35$). También se produce un efecto significativo del *tipo de prueba*, $F(1, 51) = 125.60$, $MCE = 198.72$, $p < .05$, $\eta_p^2 = 0.71$. Los ítems que van a recibir prueba de recuerdo ($M=57$, $DT=37$) obtienen estimaciones superiores a los que van a recibir prueba de reconocimiento ($M=41$, $DT=35$). El hecho de que los ítems que van a recibir prueba de recuerdo son presentados dos veces durante la fase de estudio, parece indicar que no solo aumenta el rendimiento de la prueba, sino que también aumenta la magnitud de estos juicios. Por último, replicando los resultados del Experimento 1, no se produce interacción entre las variables *tipo de juicio* y *tipo de prueba*, $F < 1$. El efecto de los juicios es muy similar en las dos pruebas.

DISTRIBUCIÓN DE LOS JUICIOS DE APRENDIZAJE

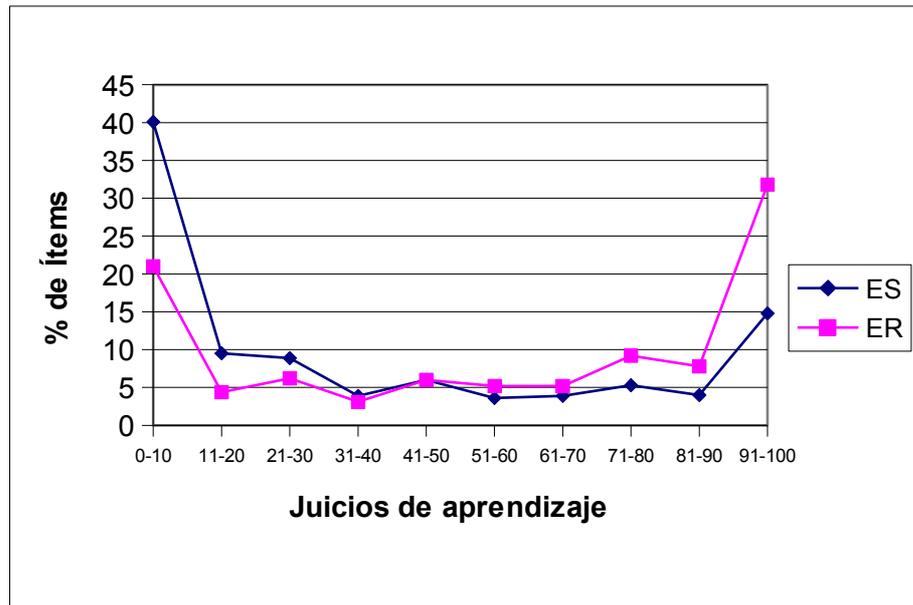


Gráfico 2.1. Distribución de los juicios de aprendizaje. Porcentaje de respuestas para cada nivel de juicio (ES: estímulo solo. ER: estímulo más respuestas).

Como podemos observar en el Gráfico 2.1, la distribución de los juicios cambia en función del estímulo presentado, siendo estas diferencias significativas: $\chi^2(9) = 17.38$ $p < .05$. La principal diferencia con respecto al experimento anterior, es que la distribución de los juicios ES pasa a ser casi bimodal. Este hecho, parece reflejar la contribución de presentar dos veces algunos ítems en este experimento.

De forma similar al Experimento 1, los juicios ES muestran una distribución asimétrica positiva. Sin embargo, los juicios ER cambian ligeramente con respecto al experimento anterior, pasando de una distribución bimodal a una distribución algo asimétrica negativa.

PRECISIÓN DE LOS JUICIOS DE APRENDIZAJE: RESOLUCIÓN

Para cada participante se calcularon 4 correlaciones gamma entre la magnitud de los juicios de aprendizaje y el rendimiento en las pruebas. El promedio de las

correlaciones gamma para cada condición puede verse en la Tabla 2.2. Una vez más realizamos aquí un análisis como el del Experimento 1, cuyos resultados pueden verse completos en el Apéndice 2.3¹⁷.

Se produce un efecto significativo del *tipo de prueba*, $F(1, 37) = 31.53$, $MCE = 0.10$, $p < .05$, $\eta_p^2 = 0.46$. Este efecto replica los resultados del Experimento 1, las estimaciones son más acertadas en la prueba de recuerdo ($M = 0.77$, $DT = 0.30$) que en la prueba de reconocimiento ($M = 0.48$, $DT = 0.28$).

	RD	RN
ES	0.78 (0.31)	0.33 (0.32)
ER	0.77 (0.29)	0.63 (0.25)

Tabla 2.2. Promedio de la correlación gamma y desviación típica (ES: estímulo solo, ER: estímulo más respuestas, RD: recuerdo, RN: reconocimiento).

También se produce un efecto significativo del *tipo de juicio*, $F(1, 37) = 8.41$, $MCE = 0.09$, $p < .05$, $\eta_p^2 = 0.18$. La resolución es superior en la condición ER ($M = 0.70$, $DT = 0.27$) que en la condición ES ($M = 0.55$, $DT = 0.32$). Ahora bien, como se ve en la Tabla 2.2, este efecto es engañoso si no se tiene en cuenta la interacción.

De la misma forma que en el Experimento 1, se produce interacción significativa entre las variables *tipo de juicio* y *tipo de prueba*, $F(1, 37) = 20.50$, $MCE = 0.05$, $p < .05$, $\eta_p^2 = 0.36$. Como en el experimento anterior, el efecto del *tipo de juicio* se da sólo en la prueba de reconocimiento, donde las diferencias ES-ER son significativas, $t(49) = 3.36$, $p < .05$. Por el contrario, en la prueba de recuerdo las diferencias ES-ER no son significativas, $t(40) = 0.25$, $p > .10$, como queda claro en la tabla.

¹⁷ En este análisis se producen las siguientes gammas indeterminadas: ESRD=5, ESRN=1, ERRD=8 y ERRN=2.

PRECISIÓN DE LOS JUICIOS DE APRENDIZAJE: AJUSTE

El grado de ajuste entre las predicciones de los participantes y su rendimiento se calculó de la misma forma que en el Experimento 1. En la prueba de reconocimiento se utilizó de nuevo la fórmula de corrección por adivinación como medida del reconocimiento neto. Los resultados se muestran en el Gráfico 2.2, en el que la diagonal muestra la calibración perfecta.

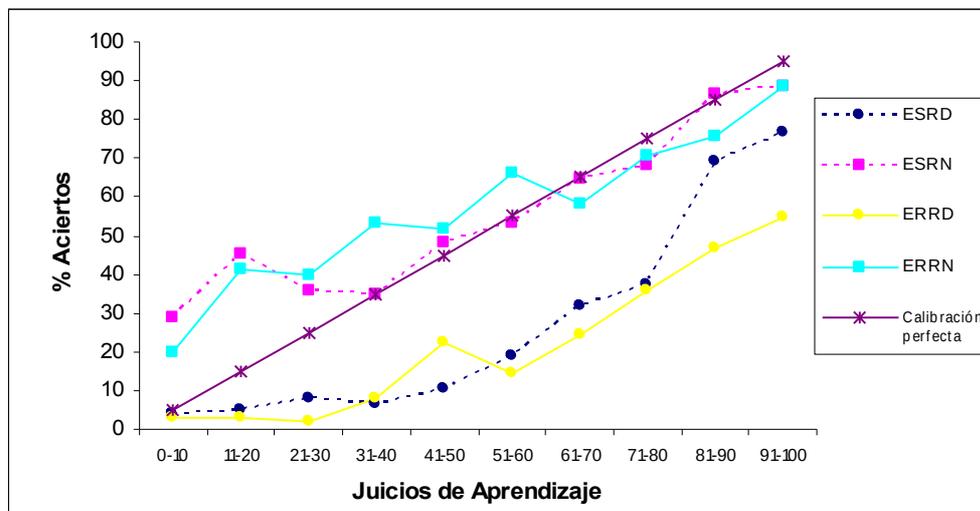


Gráfico 2.2. Curvas de calibración para cada condición experimental (ES: estímulo solo, ER: estímulo más respuestas, RD: recuerdo. RN: reconocimiento).

Para las cuatro curvas de calibración de las condiciones experimentales se ha calculado sus correspondientes rectas de regresión, que se muestran en la Tabla 2.3.

	RD	RN
ES	$y=-13.66+0.81x$	$y=22.64+0.66x$
ER	$y=-8.29+0.59x$	$y=24.85+0.63x$

Tabla 2.3. Rectas de regresión lineal. Calibración perfecta: $y=0+1x$

Los ajustes son significativos en las cuatro rectas (Ver Apéndice 2.4), mostrando de nuevo correlaciones elevadas, como podemos observar en la Tabla 2.4.

De la misma forma que en el experimento anterior, se comparan las rectas de regresión mediante el estadístico ZPF. Los ajustes no muestran diferencias significativas. Se replican por tanto, los resultados del Experimento 1.

	RD	RN
ES	0.83	0.89
ER	0.90	0.92

Tabla 2.4. Coeficiente de determinación R^2 de las rectas de regresión.

DISCUSIÓN

De acuerdo con el objetivo que nos planteábamos en este experimento, se produce un aumento del rendimiento en la prueba de recuerdo. Y hemos conseguido, además, que el rendimiento en la prueba de reconocimiento no sufra alteraciones importantes: pasa de un promedio del 57.86% de respuestas acertadas en el Experimento 1 al 61.75% de respuestas acertadas en el Experimento 2.

Aunque la repetición durante la fase de estudio de los ítems que van a recibir prueba de recuerdo provoca un aumento en la magnitud de los juicios de estos ítems sobre la de los que recibirán prueba de reconocimiento (véase p. ej., Lovelace, 1984), este hecho no provoca alteraciones sobre el efecto de la variable *tipo de juicio*, que es similar en los dos experimentos. En lo esencial, respecto de la magnitud de los juicios y su distribución se replican los resultados básicos del Experimento 1, algo que apoya de nuevo la idea de que los participantes utilizaron diferentes procesamientos ante cada contexto de JA.

De acuerdo con las predicciones de la SAT, para la resolución de las estimaciones se produce interacción entre las variables *tipo de juicio* y *tipo de*

prueba. Debemos esperar más precisión predictiva relativa del recuerdo en contextos ES y más precisión predictiva relativa del reconocimiento en contextos ER. Sin embargo, como en el experimento anterior, la interacción localiza en el efecto de los diferentes juicios sobre la prueba de reconocimiento. Aunque en el Experimento 2 aumenta el nivel de ejecución de la prueba de recuerdo, esto no afecta a la resolución de los juicios, cuya precisión en recuerdo es similar en ambas condiciones de juicios en los dos experimentos. Por su parte, el efecto de los juicios sobre la prueba de reconocimiento se muestra muy consistente y se replican con claridad los resultados del Experimento 1.

9.3- EXPERIMENTO 3.

Los resultados del Experimento 2 replican de nuevo la interacción entre las variables *tipo de juicio* y *tipo de prueba*, cumpliéndose con las predicciones de la SAT, al menos en reconocimiento. Según nuestra hipótesis, en Weaver y Kelemen (2003) no se produce esta interacción, porque los participantes basan sus decisiones, tanto en los juicios de la condición C4 como en la prueba de reconocimiento, en el componente de recuperación.

Uno de los objetivos del experimento anterior fue comprobar si la resolución de los juicios respecto a la prueba de recuerdo es inferior con juicios ante ER que con juicios ante ES. Observamos que la fuerza de la interacción en los Experimentos 1 y 2 recae sobre la prueba de reconocimiento. Por este motivo, en este tercer experimento nuestro objetivo principal será probar la hipótesis SAT en reconocimiento, para explicar la menor resolución ante juicios ES que ante juicios ER en esta prueba. Para ello, intervendremos en la medida de nuestras posibilidades sobre los componentes de recuperación y familiaridad en reconocimiento (Jacoby y Dallas, 1981; Mandler, 1980; Tulving, 1985; Yonelinas, 2007; Whittlesea, 2003). Pretendemos alterar el peso de uno y otro componentes, con el fin de demostrar que nuestro contexto ER de predicción del reconocimiento, cuyos beneficios hemos comprobado en los dos experimentos anteriores, pierde eficiencia en la predicción de reconocimiento cuando en esta tarea los procesos de recuperación son dominantes. En otras palabras, cuando la familiaridad sea el componente dominante en la tarea de reconocimiento la interacción entre el contexto de juicio y el tipo de prueba debe aparecer, como ha ocurrido en los Experimentos 1 y 2. Sin embargo, cuando el componente dominante en reconocimiento sea la recuperación esta interacción no debe aparecer, como ocurre en la investigación de Weaver y Kelemen (2003).

En el Experimento 3 vamos a trabajar con la variable *tipo de distractor*. Las palabras distractoras son las que aparecen junto a la palabra objetivo en la prueba de reconocimiento. Vamos a emplear una manipulación entre-sujetos con dos grupos experimentales: grupo de distractores internos (DI) que está formado por palabras que aparecen anteriormente en el experimento como objetivos de la prueba y el grupo de distractores externos (DE) está formado por palabras externas al experimento, es decir, palabras nuevas que no habían aparecido anteriormente.

En el grupo distractores internos (DI) intentaremos potenciar el componente de recuperación. Para ello, todas las opciones en la prueba de reconocimiento serán objetivos de otros ítems estudiados. Con esto, conseguimos que todas las opciones estén igualadas en familiaridad, fomentando así el uso de la recuperabilidad por parte de los participantes. Recuérdese que este *tipo de distractor* es similar al utilizado en el experimento de Weaver y Kelemen (2003).

Por el contrario, en el grupo de distractores externos (DE), intentaremos potenciar el componente de familiaridad. Para ello, todos los distractores de la prueba de reconocimiento serán palabras externas al experimento, palabras que nunca antes han aparecido. De este modo pretendemos provocar el uso del contraste de familiaridad entre objetivo y distractores como propiedad discriminativa por parte de los participantes.

Si nuestra hipótesis es correcta, esperamos que: en el grupo distractores internos (DI) la interacción entre las variables *tipo de juicio* y *tipo de prueba* desaparezca; mientras que, por el contrario, en el grupo de distractores externos (DE), la interacción observada en los Experimentos 1 y 2 debe ser replicada.

MÉTODO

PARTICIPANTES

Participaron 66 estudiantes universitarios de la Facultad de Psicología de la UNED, 34 alumnos en la DI (29 mujeres y 5 hombres), con una edad comprendida entre los 23 y 66 años ($M= 41.06$, $DT= 10.29$) y 32 alumnos en DE (25 mujeres y 7 hombres), con una edad comprendida entre los 26 y 58 años ($M= 38.63$, $DT= 8.51$), para obtener créditos en la asignatura *Psicología de la Memoria*.

MATERIALES Y APARATOS

Los materiales y aparatos son los mismos que en el Experimento 2. Los distractores internos (DI) son seleccionados al azar entre los términos repuestas. Puesto que son necesarias 120 respuestas y solo tenemos 80 posibles, algunas se repiten. Respecto a los distractores externos (DE), son seleccionadas al azar 120 palabras adicionales, la mitad de alta frecuencia y la otra mitad de baja frecuencia.

PROCEDIMIENTO Y DISEÑO

En las tareas de adquisición y juicios de aprendizaje no se introduce ningún cambio respecto a los experimentos anteriores. Para las *Pruebas de Memoria* se introducen los siguientes cambios:

Grupo distractores internos (DI). En la prueba de reconocimiento, de las 4 alternativas de respuesta, 1 será el objetivo y los 3 distractores serán también objetivos seleccionados al azar de otros ítems presentados.

Grupo distractores externos (DE). En la prueba de reconocimiento, de las 4 alternativas de respuesta, 1 será el objetivo y los 3 distractores serán palabras externas al experimento.

Todos los demás detalles del procedimiento fueron como los descritos en el Experimento 2.

RESULTADOS

RENDIMIENTO EN LAS PRUEBAS DE MEMORIA

Podemos ver el promedio del rendimiento en las pruebas en la Tabla 3.1. Realizamos sobre estos datos un ANOVA mixto con la variable *tipo de distractor* como factor entre-sujetos y las variables tipo de juicio y tipo de prueba como factores intrasujetos. Puede verse el detalle de los resultados en el Apéndice 3.1.

Se produce un efecto significativo de la variable *tipo de distractor*, $F(1, 64) = 7.22$, $MCE = 4204.32$, $p < .05$, $\eta_p^2 = 0.10$. El rendimiento es superior en DE ($M=57$, $DT=36$) que en DI ($M=42$, $DT=39$). Además, la interacción entre las variables *tipo de distractor* y *tipo de prueba* es significativa, $F(1, 64) = 12.59$, $MCE = 438.35$, $p < .05$, $\eta_p^2 = 0.16$, debido a que el efecto de la variable *tipo de distractor* es superior en reconocimiento que en recuerdo, como podemos observar en la Tabla 3.1. Este resultado tiene especial interés para nosotros, ya que indica que la habitual diferencia entre recuerdo y reconocimiento es mayor cuando se usan distractores externos. Parece que el uso de palabras externas en la prueba de reconocimiento proporciona un potente contraste de familiaridad entre objetivo y distractores como clave de reconocimiento.

	Distractores Internos		Distractores Externos	
	RD	RN	RD	RN
ES	31 (46)	47 (27)	40 (49)	66 (21)
ER	40 (49)	49 (32)	49 (50)	73 (22)

Tabla 3.1. Promedio y desviación típica de los porcentajes de aciertos en las pruebas de memoria (ES: estímulo solo, ER: estímulo más respuestas, RD: recuerdo, RN: reconocimiento).

Replicando los resultados encontrados en los Experimentos 1 y 2, se produce un efecto significativo del *tipo de juicio*, $F(1, 64) = 23.24$, $MCE = 262.53$, $p < .05$, $\eta_p^2 = 0.27$, y del *tipo de prueba*, $F(1, 64) = 103.01$, $MCE = 438.35$, $p < .05$, $\eta_p^2 = 0.62$. El rendimiento en las pruebas es superior cuando hemos presentado ER en los juicios (DI: $M=44$, $DT=40$ y DE: $M=61$, $DT=36$) que cuando hemos presentado ES (DI: $M=39$, $DT=36$ y DE: $M=53$, $DT=35$). En la prueba de reconocimiento (DI: $M=48$, $DT=29$ y DE: $M=69$, $DT=21$) el rendimiento es más elevado que en la prueba de recuerdo (DI: $M=35$, $DT=47$ y DE: $M=44$, $DT=49$). Por último, como en el Experimento 1, la interacción entre las variables *tipo de juicio* y *tipo de prueba* estuvo próxima a la significación, $F(1, 64) = 3.58$, $MCE = 210.62$, $p < .10$, $\eta_p^2 = 0.05$, ya que el efecto del *tipo de juicio* fue algo más grande en recuerdo que en reconocimiento. Finalmente, la interacción de segundo orden que implica a los tres factores (*tipo de juicio*, *tipo de prueba* y *tipo de distractores*) no resultó estadísticamente significativa ($F < 1$).

MAGNITUD DE LOS JUICIOS DE APRENDIZAJE

Una vez más, calculamos el promedio de la magnitud de los juicios de aprendizaje de la misma forma que en los anteriores experimentos, como se muestra en la Tabla 3.2. Pueden verse los resultados completos en el Apéndice 3.1.

No se producen efecto principal significativo de la variable *tipo de distractor*, $F(1, 64) = 0.08$, $MCE = 1969.25$, $p > .10$, $\eta_p^2 = 0.00$. Es importante destacar este aspecto, que nos apunta que la tarea de juicios es similar en ambas condiciones.

	Distractores Internos		Distractores Externos	
	RD	RN	RD	RN
ES	52 (39)	36 (35)	53 (40)	31 (34)
ER	73 (35)	58 (38)	79 (32)	60 (38)

Tabla 3.2. Promedio y desviación típica de la magnitud de los juicios (ES: estímulo solo, ER: estímulo más respuestas, RD: recuerdo, RN: reconocimiento).

Se produce un efecto principal significativo del *tipo de juicio*, $F(1, 64) = 214.04$, $MCE = 360.66$, $p < .05$, $\eta_p^2 = 0.77$; así como del *tipo de prueba*, $F(1, 64) = 274.85$, $MCE = 153.8$, $p < .05$, $\eta_p^2 = 0.81$; y, por último, no se produce la interacción entre las variables *tipo de juicio* y *tipo de prueba*, $F(1, 64) = 1.38$, $MCE = 116.97$, $p > .10$, $\eta_p^2 = 0.02$, como tampoco se produjo la de segundo orden que implica a los tres factores, $F(1, 64) = 0.31$, $MCE = 116.97$, $p > .10$, $\eta_p^2 = 0.00$. En conjunto, los resultados relativos a la magnitud de los juicios de aprendizaje obtenidos en este experimento replican los hallazgos de los Experimentos 1 y 2 y señalan, una vez más, que la tarea de juicios de aprendizaje de ambas condiciones no puede verse afectada por características de una prueba de memoria posterior que los individuos desconocen.

DISTRIBUCIÓN DE LOS JUICIOS DE APRENDIZAJE

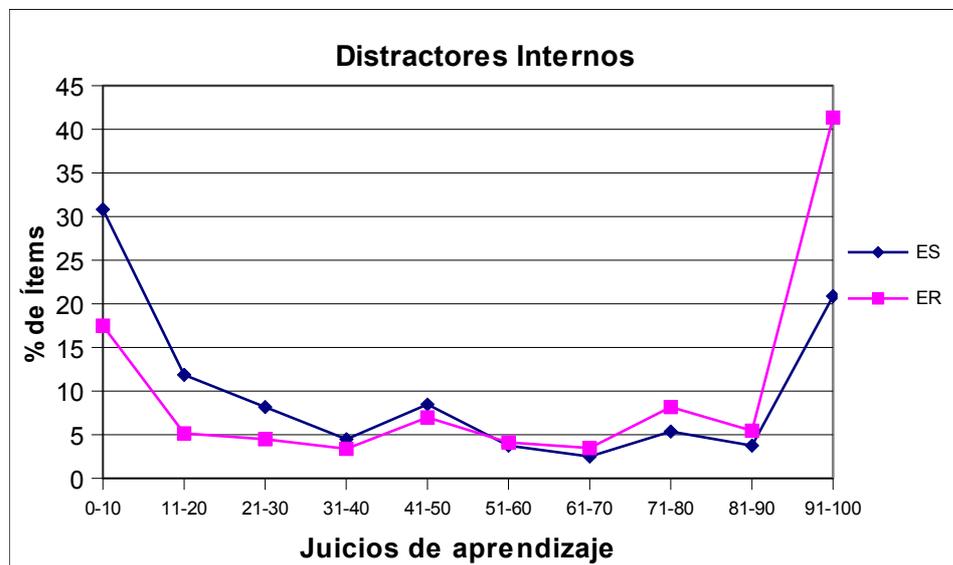


Gráfico 3.1. Distribución de los juicios de aprendizaje para DI. Porcentaje de respuestas para cada nivel de juicio (ES: estímulo solo. ER: estímulo más respuestas).

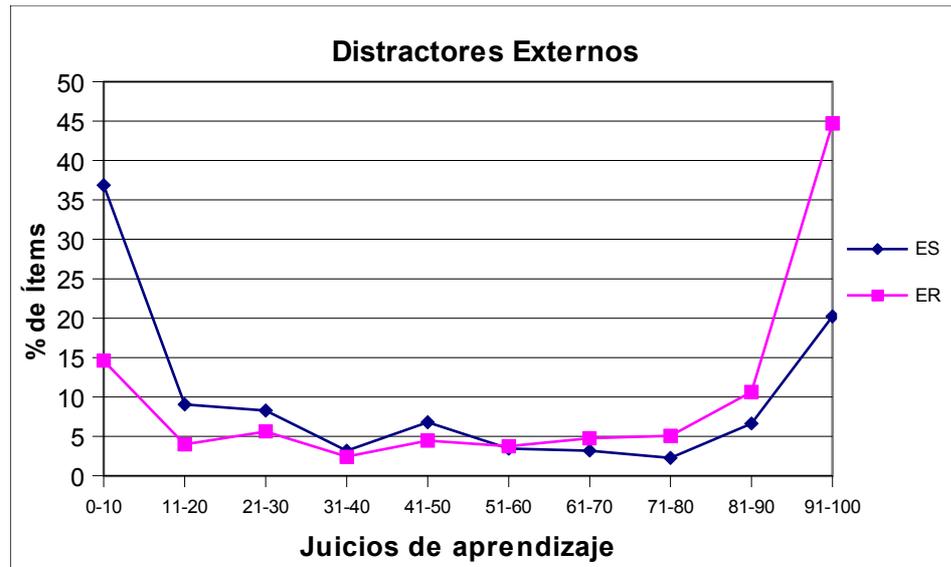


Gráfico 3.2. Distribución de los juicios de aprendizaje para DE. Porcentaje de respuestas para cada nivel de juicio (ES: estímulo solo. ER: estímulo más respuestas).

Como podemos observar en los Gráficos 3.1 y 3.2, en ambos grupos se da una distribución de los juicios muy similar. Los juicios ES muestran una distribución asimétrica positiva, mientras que, los juicios ER muestran una distribución asimétrica negativa, llegando a ser casi bimodal en ambos casos. La similitud de resultados entre ambas condiciones, una vez más indica, por la razones que acabamos de señalar en el punto anterior, que los participantes emiten sus juicios de la misma forma en las dos condiciones.

Por otro lado, la distribución de los juicios cambia en función del estímulo presentado, siendo estas diferencias significativas en DE: $\chi^2(9) = 24.24$ $p < .05$ y aproximándose a la significación en DI: $\chi^2(9) = 15.44$ $p < 0.10$.

PRECISIÓN DE LOS JUICIOS DE APRENDIZAJE: RESOLUCIÓN

La correlación gamma se calculó de la misma forma que en los experimentos anteriores. El promedio de dichas correlaciones en cada *tipo de distractor* puede verse en la Tabla 3.3. Un ANOVA mixto de tres factores (2 x (2x2xS)), considerando el *tipo de distractor* como factor entre grupos y el *tipo de juicio* y *tipo de prueba*

como factores intrasujetos, mostró significativo el factor *tipo de prueba*, $F(1, 55) = 31.15$, $MCE=.18$, $p<.05$, $\eta_p^2 = .36$, así como la interacción de segundo orden que implica a los tres factores, $F(1, 55) = 5.08$, $MCE=.118$, $p<.05$, $\eta_p^2 = .08$ (en el Apéndice 3.3.1 puede verse la información detallada sobre este análisis). El resultado interesante para nosotros es esta interacción, que, como se observa en la Tabla 3.3, parece deberse a que existe una diferencia fundamental entre los dos grupos en los efectos de los factores *tipo de juicio* y *tipo de prueba*. Para descomponer esta interacción y verificar si se cumplen los resultados esperados en este experimento, se analizaron por separado los grupos DI y DE, utilizando el mismo procedimiento de análisis¹⁸ que en el Experimento 2. Pueden verse los resultados completos en el Apéndice 3.3.2.

	Distractores Internos		Distractores Externos	
	RD	RN	RD	RN
ES	0.80 (0.26)	0.54 (0.29)	0.81 (0.33)	0.36 (0.47)
ER	0.76 (0.32)	0.38 (0.51)	0.72 (0.43)	0.57 (0.39)

Tabla 3.3. Promedio de la correlación gamma y desviación típica para cada *tipo de distractor* (ES: estímulo solo, ER: estímulo más respuestas, RD: recuerdo, RN: reconocimiento).

Un primer resultado importante es que en DE, con distractores externos en la prueba de reconocimiento, replicando los hallazgos encontrados en los Experimentos 1 y 2, se produce interacción entre las variables *tipo de juicio* y *tipo de prueba*, $F(1, 26) = 4.39$, $MCE = 0.14$, $p<.05$, $\eta_p^2 = 0.14$, una vez más debido a que los JA (Juicios de Aprendizaje) ante ER presentaron, como predictores de reconocimiento, más resolución que los realizados ante ES, $t(26)=2.07$, $p<.05$, mientras que en la prueba de recuerdo esta relación parece invertirse, aunque sin alcanzar significación estadística, $t(26)=0.80$, $p>.10$. Como hemos dicho, este resultado, que claramente

¹⁸ En este análisis se producen las siguientes gammas indeterminadas. DI: ESRD=2, ESRN=1, ERRD=2 y ERRN=1. DE: ESRD=1, ESRN=1, ERRD=1 y ERRN=4.

difiere de los informados por Weaver y Kelemen (2003), suponen una nueva replicación de lo obtenido en los dos experimentos anteriores.

Pero el resultado más importante de este Experimento, quizá sea, que en DI, en la que los distractores en reconocimiento son internos al experimento, la interacción entre las variables *tipo de juicio* y *tipo de prueba* desaparece, $F(1, 29) = 0.89$, $MCE = 0.10$, $p > .10$, $\eta_p^2 = 0.03$. Como podemos observar en la Tabla 3.3, tanto respecto a la prueba de recuerdo como a la de reconocimiento, las resoluciones de los JA son más elevadas ante ES que ante ER. La importancia de este resultado, obtenido con distractores intraexperimentales en la prueba de reconocimiento, radica en que, en contraste con lo obtenido en nuestros dos experimentos anteriores, replica en esencia los resultados informados por Weaver y Kelemen (2003), cuyos distractores también eran palabras usadas en la fase de estudio del procedimiento.

PRECISIÓN DE LOS JUICIOS DE APRENDIZAJE: AJUSTE

De nuevo calculamos el grado de ajuste como en los experimentos anteriores.

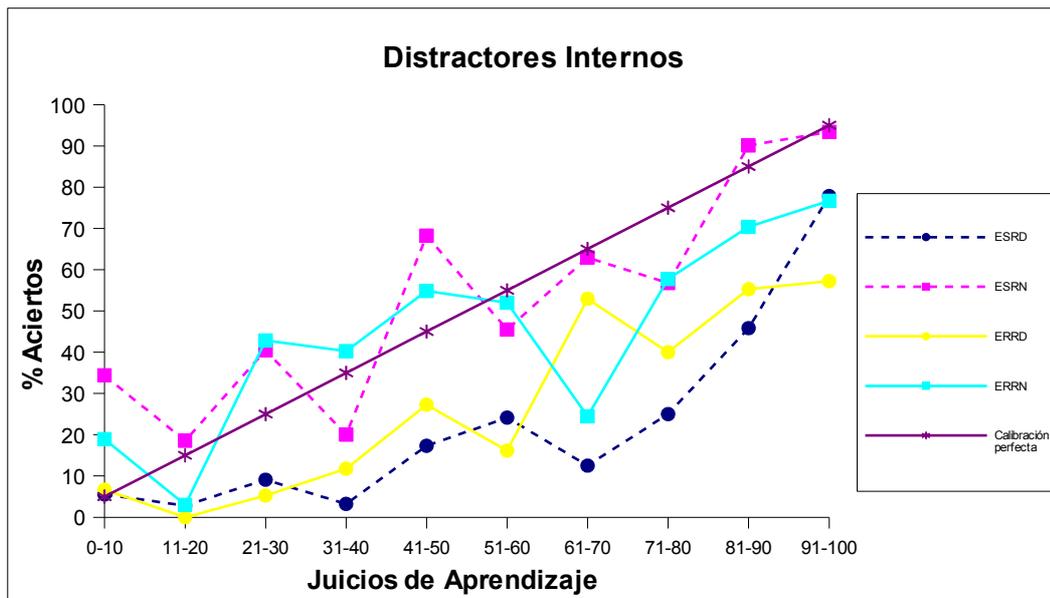


Gráfico 3.3. Curvas de calibración de DI (ES: estímulo solo, ER: estímulo más respuestas, RD: recuerdo, RN: reconocimiento).

Podemos ver los resultados en los Gráficos 3.3 y 3.4, donde la diagonal muestra la calibración perfecta.

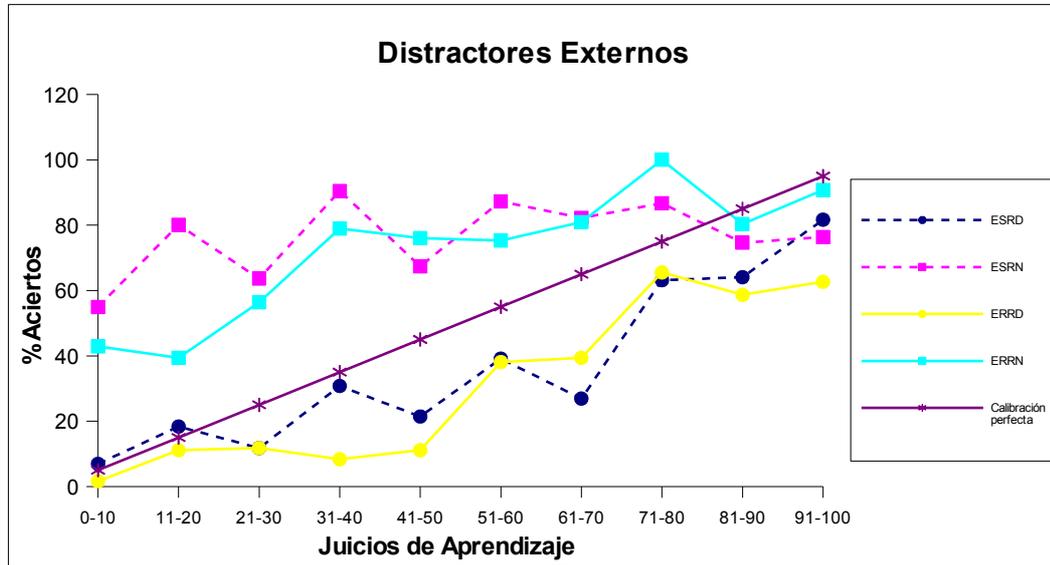


Gráfico 3.4. Curvas de calibración de DE (ES: estímulo solo, ER: estímulo más respuestas, RD: recuerdo. RN: reconocimiento).

En la Tabla 3.4 se muestran las rectas de regresión correspondientes a las cuatro curvas de calibración de cada *tipo de distractor*.

	Distractores Internos		Distractores Externos	
	RD	RN	RD	RN
ES	$y = -9,98 + 0,65x$	$y = 16,06 + 0,74x$	$y = -1,63 + 0,76x$	$y = 68,36 + 0,16x$
ER	$y = -6,93 + 0,68x$	$y = 13,34 + 0,61x$	$y = -7,68 + 0,77x$	$y = 43,58 + 0,57x$

Tabla 3.4. Rectas de regresión lineal para cada *tipo de distractor*. (Calibración perfecta: $y=0+1x$).

Todas las rectas muestran ajustes significativos excepto la recta ESRN-DE (ver Apéndice 3.4), cuyo coeficiente de determinación es muy inferior al de las demás. Como podemos observar en la Tabla 3.5, el ajuste entre juicios y rendimiento en la prueba de reconocimiento de DE es superior ante ER que ante ES, mientras que, en DI es superior ante ES que ante ER.

	Distractores Internos		Distractores Externos	
	RD	RN	RD	RN
ES	0.69	0.73	0.84	0.18
ER	0.85	0.65	0.87	0.76

Tabla 3.5. Coeficiente de determinación R^2 de las rectas de regresión en cada *tipo de distractor*.

Cuando comparamos las rectas de regresión mediante el estadístico ZPF, solo observamos diferencias significativas entre las rectas: ESRD-ESRN-DE, $ZPF(N=10)=2.40$, $p<.05$ y ESRN-ERRN-DE, $ZPF(N=10)=2.16$, $p<.05$ (ver Apéndice 3.4). Este patrón de resultados es plenamente compatible con la hipótesis SAT, ya que muestra cómo los JA emitidos en el contexto ES presentan una mayor calibración absoluta respecto al recuerdo que al reconocimiento. Ahora bien, respecto a la prueba de reconocimiento, el ajuste de los JA es superior ante ER que ante ES.

En resumen, los JA emitidos en el contexto ES presentan un claro desajuste de calibración absoluta como predictores de reconocimiento con distractores externos; mientras que los emitidos en el contexto ER presentan un ajuste próximo al resto de las condiciones de nuestro experimento. Enseguida volveremos sobre las implicaciones teóricas de este resultado.

DISCUSIÓN

El rendimiento en reconocimiento es superior en DE que en DI. Esto puede deberse, a que en la prueba de reconocimiento de DE, a diferencia de lo que ocurre en DI, el nivel de familiaridad de la palabra objetivo es presumiblemente superior al de los distractores. Esto contribuye a que los participantes puedan ayudarse también del componente de familiaridad, aumentando así su rendimiento en la prueba. Por el contrario, en la prueba de reconocimiento de DI, igualamos la familiaridad del

objetivo y los distractores, empujando a la persona a basar el reconocimiento casi exclusivamente en el componente de recuperación. Suponemos que con una menor contribución del componente de familiaridad, su rendimiento disminuye.

Respecto a la magnitud de los juicios y su distribución, hemos visto que, como esperábamos, en lo esencial se replican los resultados obtenidos en los Experimentos 1 y 2. Como la diferencia entre las dos condiciones aparece una vez emitidos los juicios, es evidente que no hay motivos para esperar diferencias entre los juicios de un *tipo de distractor* y otro en lo relativo a su magnitud y distribución, más allá de las atribuibles a las naturales diferencias individuales.

La resolución de la prueba de recuerdo, muestra resultados similares en ambas condiciones, replicando los resultados de los experimentos anteriores.

Por su parte, la resolución de la predicción del reconocimiento muestra los resultados esperados por nuestra hipótesis. En DI, la neutralización del componente de familiaridad en la prueba de memoria obliga a los participantes a basar su reconocimiento esencialmente en el componente de recuperación. Por este motivo, al aproximar el procesamiento en la prueba de reconocimiento al típico de una prueba de recuerdo, la clave de juicio ES se muestra más efectiva para la resolución de los JA que la clave ER. Nótese que esta diferencia también debería apreciarse respecto a la prueba de recuerdo, aunque una resolución ciertamente alta de JA ante ER podría haber neutralizado la diferencia por un efecto de techo. Por el contrario, en DE, en la que una vez más usamos distractores externos en reconocimiento, permitiendo así jugar un papel más relevante al componente de familiaridad, la interacción observada en nuestros Experimentos 1 y 2 se replica, confirmando así que respecto a un reconocimiento con un fuerte componente de familiaridad, la clave ER es más efectiva, como se esperaba desde los supuestos de la SAT.

En cuanto al ajuste o *calibración absoluta* entre juicios y rendimiento, el resultado más relevante parece ser el especial deterioro del ajuste de los juicios ante ES como predictores del rendimiento en reconocimiento en DE. La magnitud del desajuste (.18) se aprecia mejor si lo comparamos con el alto ajuste alcanzado por estos mismos participantes y las mismas claves de juicio en JA como predictores del recuerdo (.84). Una vez más, parece que el fuerte componente de familiaridad en la tarea de reconocimiento, cuando los distractores son externos, hace que JA emitidos ante el término estímulo del par asociado resulten erráticos como predictores del rendimiento en tales pruebas de memoria. Este hecho es perfectamente congruente con la idea de que los JA emitidos ante este tipo de claves se basan en buena medida en la recuperabilidad de la respuesta asociada, como vimos en la introducción de esta tesis (Dunlosky y Nelson, 1994, 1997; Nelson, Narens y Dunlosky, 2004). Por el contrario, parece que nuestra prueba de reconocimiento para estos participantes tiene presumiblemente un elevado componente de familiaridad. En apoyo de tal interpretación vemos cómo el desajuste se convierte en un nivel de ajuste más que aceptable (.76) precisamente cuando los JA se emiten en un contexto ER (un distractor interno y dos externos), en el que el propio juicio puede beneficiarse de la familiaridad diferencial de la alternativa correcta.

10.- DISCUSIÓN GENERAL Y CONCLUSIONES.

Hemos investigado la capacidad de las personas para predecir su rendimiento en pruebas de recuerdo y reconocimiento. Y lo hemos hecho desde los postulados de la hipótesis de la *Supervisión Apropriada para la Transferencia* (SAT). La SAT nos ofrece una explicación teórica del acierto o desacierto con que las personas realizan los juicios metamnemónicos (Begg y otros, 1989; Ruiz, 2004; Thiede y Dunlosky, 1994). Se basa para ello en destacar la importancia de la interacción entre los procesos de codificación y recuperación en el rendimiento memorístico, que fue señalada originalmente por la hipótesis del *Procesamiento Apropriado para la Transferencia* (Craik, 2002; Morris y otros, 1977). De acuerdo con esta, la ejecución de memoria se mejora cuando el procesamiento que usamos durante la codificación de la información se asemeja al procesamiento utilizado durante la recuperación de la misma. Pues bien, ya hemos visto que, como una extensión de esta hipótesis desarrollada desde el ámbito de la memoria hacia el ámbito de la metamemoria, la SAT nos dice que la precisión de los juicios metamnemónicos depende de la similitud entre el procesamiento necesario para la elaboración de los juicios, de un lado, y el procesamiento necesario para la propia prueba de recuperación, del otro.

Hemos visto que, aunque el trabajo de Weaver y Kelemen (2003; en especial sus condiciones C1 y C4) no parece respaldar la hipótesis SAT, hay que plantearse ante él la siguiente pregunta: ¿por qué su condición C4 de juicios de aprendizaje no es más predictiva del reconocimiento posterior que su condición C1, existiendo una correspondencia aparentemente exacta entre contextos y procesamientos en la primera y no en la segunda?

Una pista para responder esta pregunta la podemos encontrar en el hecho de

que, como puede observarse en su propio trabajo, los juicios de aprendizaje emitidos en el contexto de la condición C4, cuya apariencia formal se asemeja a una prueba de reconocimiento, son tan predictivos de recuerdo como los juicios emitidos en la condición C1, que se produce en un contexto análogo al de una prueba de recuerdo. Por tanto, cabe la posibilidad de que las condiciones estímulares para la realización de los juicios de aprendizaje en la condición C4 generen un procesamiento similar al que suscitan las condiciones estímulares de C1, a pesar de que los contextos de estas dos condiciones no sean nominalmente equiparables.

Como vimos en la introducción de esta tesis, Begg y otros (1989, Experimento 4) mostraron que son las condiciones estímulares (factor contextual) las que generan un tipo u otro de procesamiento, independientemente del tipo de prueba que esperan los participantes (factor estratégico). Vimos también que estos hallazgos fueron en lo esencial replicados por Thiede y Dunlosky (1994).

Por otro lado, hemos visto que en una prueba de reconocimiento asociativo la persona puede hacer uso de la recuperación y de la familiaridad de la información objetivo (p. ej., Castel y Craik, 2003; Lee y otros, 2013; Buchler y otros, 2008; Mandler, 1980, 1991; Parks y Yonelinas, 2015; Verde, 2004; Yonelinas y otros, 2010). El predominio de un tipo u otro de información en la que basar sus decisiones dependerá en gran medida de factores situacionales (factor contextual). Por ejemplo, cuando los distractores son diferentes entre sí y respecto al objetivo en diversas propiedades no determinadas a priori (randomly different, en términos de los autores) la mejor estrategia es basar el reconocimiento en la familiaridad (Malmberg, Zeelenberg y Shiffrin, 2004). Mientras que, cuando hay una gran similitud entre las propiedades fundamentales de los objetivos y la de los distractores, es más eficaz basar el reconocimiento en la recuperabilidad (Malmberg, Holden y Shiffrin, 2004).

A pesar de ello, en ocasiones los individuos no son capaces de controlar el juego combinado de la recuperación y la familiaridad y se basan exclusivamente en uno de los mecanismos. Así ocurre cuando, el tiempo disponible para emitir la respuesta es limitado, algo que obliga a los individuos a basarse en la familiaridad del ítem, ya que se supone que es una información de acceso más rápido (Malmberg y Xu, 2007; Xu y Malmberg, 2007; Malmberg, 2008).

Nuestra hipótesis parte de la idea de que el factor contextual de C4 activa procesos de recuperación semejantes a los de una prueba de recuerdo. En C4 los estímulos presentados como contexto de juicio no son estímulos "aleatoriamente diferentes del objetivo", sino que están igualados en familiaridad. Esto es debido a que tanto los objetivos como los distractores son palabras del experimento. De esta forma, el papel de la familiaridad queda neutralizado (Yonelinas, 1997; Malmberg, Holden y Shiffrin, 2004).

Pues bien, el propósito de nuestros Experimentos 1 y 2 fue potenciar el papel de la familiaridad. Para ello en los juicios ER y en la prueba de reconocimiento introdujimos como distractores palabras externas al experimento. Así, conseguimos que la familiaridad pase a ser una información relevante, y por lo tanto, pueda ser utilizada como una clave *mnemónica* en los juicios, por usar la terminología de la aproximación basada en el uso de las claves de Koriat (1997)

Los resultados del Experimento 1 revelan que los participantes usaron efectivamente la familiaridad como una fuente de información para sus estimaciones. Las condiciones estimulares de los juicios generaron diferentes procesamientos de la información. Como lo demuestra el cambio en la magnitud de los juicios y la distribución de los mismos. Evidenciando que ante los juicios ES predominan los procesos de recuperación. Mientras que, ante los juicios ER, además de los procesos

de recuperación, los participantes también hacen uso de la familiaridad como clave mnemónica. Todo esto se ve respaldado por un resultado crucial de este experimento: la interacción entre juicios y pruebas de memoria para la precisión de las estimaciones. Estos hallazgos respaldan la hipótesis SAT y además apoyan nuestra interpretación de los resultados obtenidos en su condición C4 por Weaver y Kelemen (2003): estaban basados principalmente en procesos de recuperación. Estos resultados se replicaron en los Experimentos 2 y 3.

La finalidad del Experimento 3 fue actuar en lo posible sobre los procesos de recuperación y familiaridad. Nuestro objetivo en el grupo de distractores internos (DI) fue revertir los hallazgos encontrados en los Experimentos 1 y 2. Para ello neutralizamos el papel de la familiaridad utilizando como distractores solo palabras internas de la prueba. De esta forma, igualamos el contexto de juicios y pruebas a las de Weaver y Kelemen (2003). Por su parte, el objetivo del grupo de distractores externos (DE) fue replicar los resultados de los experimentos anteriores. Para ello potenciamos el papel de la familiaridad usando solo palabras externas a la prueba como distractores.

Los resultados del Experimento 3 confirman que es posible alterar la resolución de los juicios de aprendizaje, es decir, su capacidad predictiva, simplemente actuando sobre los procesos de recuperación y familiaridad en la prueba de memoria. En el grupo de distractores internos (DI) se replican los resultados de Weaver y Kelemen (2003) y en la grupo de distractores externos (DE) se replican los resultados de nuestros Experimentos 1 y 2.

Los descubrimientos obtenidos en esta investigación tienen importantes implicaciones tanto teóricas como prácticas.

En el ámbito teórico de la metamemoria podemos argumentar que la precisión de los juicios metamnemónicos está en función de la similitud entre las condiciones durante los juicios y las condiciones durante las pruebas de memoria. Cuanto mayor sea la similitud de los procesamientos requeridos en juicios y pruebas mayor será la precisión de las estimaciones. Esta orientación, inicialmente propuesta por Begg y otros (1989) para los juicios de aprendizaje, ha sido más tarde avalada de una u otra forma por otros autores. Así, por ejemplo, en un trabajo reciente Soderstrom, Clark, Halamish y Bjork (2015) han adoptado una posición relativamente similar a la SAT. Partiendo del hecho ya conocido de que un simple JOL puede mejorar la recuperabilidad de la información (Dougherty, Sheck, Nelson y Narens, 2005), ellos han mostrado datos que apuntan a que este beneficio podría deberse al fortalecimiento durante el juicio de la misma vía asociativa que posteriormente va a permitir la recuperación. Nótese que la idea de la especificidad contextual y de proceso del juicio de aprendizaje como un modificador en sí mismo del trazo de memoria se entronca, a su vez, en el ámbito de investigaciones sobre la especificidad contextual de los beneficios de una prueba de memoria anterior sobre el rendimiento en otra posterior (p.ej., Karpicke, Lehman y Aue, 2014; McDaniel, Kowitz y Dunay, 1989).

Desde un punto de vista práctico, este trabajo puede tener importantes implicaciones en el campo de la educación y de todo tipo de programas de formación. Hay muchos trabajos que investigan cómo influyen los juicios metacognitivos en la determinación del tiempo de estudio (véase p. ej., Hines, Touron y Hertzog, 2009; Koriat y Nussinson, 2009; Koriat, Nussinson y Ackerman, 2014). Cuando un estudiante prepara una materia, evalúa constantemente su progreso en el aprendizaje y su estado de conocimiento de la misma, en función de los

resultados de estas autoevaluaciones más o menos explícitas, deberá tomar decisiones acerca de asuntos como si debe continuar estudiando del mismo modo, si puede finalizar el estudio o si debe cambiar de método de estudio. Por ejemplo, supongamos que un grupo de alumnos está aprendiendo vocabulario de una lengua extranjera, estudiando ítems como Maison-Casa (palabra francesa y su traducción al castellano); si una vez estudiado el ítem, el alumno espera un corto espacio de tiempo para emitir el juicio de aprendizaje del mismo (p. ej., estudiando otros ítems), la precisión de estos juicios será más elevada. Este efecto beneficioso de la demora de los juicios se produce tanto si la evaluación del aprendizaje del alumno se realiza mediante una prueba de recuerdo (Nelson y Dunlosky, 1991; Dunlosky y Nelson, 1992) como de reconocimiento (Thiede y Dunlosky, 1994; Dunlosky y Nelson, 1997). Pues bien, nuestra tesis se suma a este ámbito de contribuciones por cuanto sugiere que las autoevaluaciones no deben considerarse en sí mismas confiables o no (es decir, si tienen o no validez predictiva), sino que su validez dependerá en buena medida del grado en que susciten el mismo tipo de procesamientos que la prueba de evaluación posterior. Hemos demostrado que un mismo juicio puede ser predictivo de una prueba y no de otra, aun siendo ambas acerca del mismo material de aprendizaje. En consecuencia, está claro que si estas auto-evaluaciones o juicios son acertadas permitirán al alumno regular su tiempo de estudio de una forma eficiente. En este sentido, de los resultados de nuestra investigación se desprende la idea de que para conseguir una mayor precisión de las estimaciones, además de demorar los juicios es necesario que las condiciones en las que el estudiante toma sus juicios de aprendizaje deben ser lo más similares posibles a las condiciones que le serán requeridas en la prueba de memoria posterior.

Para finalizar, nuestros resultados muestran que los participantes utilizan el

componente de familiaridad en los juicios ER. Aunque, es evidente que también interviene el componente de recuperación. Como lo demuestra la elevada precisión de las estimaciones en recuerdo ante los juicios ER. Es difícil saber en que medida actúan y se combinan estos componentes en la formación de los juicios. Por este motivo, queda pendiente para futuras investigaciones continuar con el estudio de los componentes de recuperación y familiaridad en los juicios metacognitivos.

LISTA DE SÍMBOLOS, ABREVIATURAS Y SIGLAS.

DE: Distractores externos.

DI: Distractores internos.

DT: Desviación típica.

ER: Juicio con clave estímulo más respuestas.

ES: Juicio con clave estímulo solo.

JA. Juicios de Aprendizaje.

JFA: Juicios de facilidad de aprendizaje.

JIS: Juicios de impresión de saber.

M: Media.

RA: Reconocimiento asociativo

RI: Reconocimiento del ítem

RN: Prueba de reconocimiento.

RD: Prueba de recuerdo.

SAT: Supervisión Apropiaada para la Transferencia.

seg: Segundo

LISTA DE TABLAS Y FIGURAS Y GRÁFICOS.

Tablas.

Tabla 1.1. Promedio y desviación típica de la magnitud de los juicios y del porcentaje de aciertos en las pruebas de memoria (ES: estímulo solo, ER: estímulo más respuestas, RD: recuerdo, RN: reconocimiento). Página 37.

Tabla 1.2. Promedio de la correlación gamma y desviación típica (ES: estímulo solo, ER: estímulo más respuestas, RD: recuerdo, RN: reconocimiento). Página 39.

Tabla 1.3. Rectas de regresión lineal. Calibración perfecta: $y=0+1x$. Página 41.

Tabla 1.4. Coeficiente de determinación R^2 de las rectas de regresión. Página 41.

Tabla 2.1. Promedio y desviación típica de la magnitud de los juicios y del porcentaje de aciertos en las pruebas de memoria (ES: estímulo solo, ER: estímulo más respuestas, RD: recuerdo, RN: reconocimiento). Página 48.

Tabla 2.2. Promedio de la correlación gamma y desviación típica (ES: estímulo solo, ER: estímulo más respuestas, RD: recuerdo, RN: reconocimiento). Página 50.

Tabla 2.3. Rectas de regresión lineal. Calibración perfecta: $y=0+1x$. Página 51.

Tabla 2.4. Coeficiente de determinación R^2 de las rectas de regresión. Página 52.

Tabla 3.1. Promedio y desviación típica de los porcentajes de aciertos en las pruebas de memoria (ES: estímulo solo, ER: estímulo más respuestas, RD: recuerdo, RN: reconocimiento). Página 57.

Tabla 3.2. Promedio y desviación típica de la magnitud de los juicios (ES: estímulo solo, ER: estímulo más respuestas, RD: recuerdo, RN: reconocimiento). Página 58.

Tabla 3.3. Promedio de la correlación gamma y desviación típica para cada tipo de distractor (ES: estímulo solo, ER: estímulo más respuestas, RD: recuerdo, RN: reconocimiento). Página 61.

Tabla 3.4. Rectas de regresión lineal para cada tipo de distractor.(Calibración perfecta: $y=0+1x$). Página 63.

Tabla 3.5. Coeficiente de determinación R² de las rectas de regresión en cada tipo de distractor. Página 64.

Figuras y Gráficos.

Figura 1.1. Adaptación del esquema de metamemoria propuesto por Nelson y Narens (1990). En el centro aparecen las etapas del proceso de memoria, en la parte superior algunos ejemplos de componentes de supervisión y en la parte inferior algunos ejemplos de componentes de control. Página 6.

Gráfico 1.1. Distribución de los juicios de aprendizaje. Porcentaje de respuestas para cada nivel de juicio (ES: estímulo solo. ER: estímulo más respuestas). Página 38.

Gráfico 1.2. Curvas de calibración para cada condición experimental (ES: estímulo solo, ER: estímulo más respuestas, RD: recuerdo. RN: reconocimiento). Página 40.

Gráfico 2.1. Distribución de los juicios de aprendizaje. Porcentaje de respuestas para cada nivel de juicio (ES: estímulo solo. ER: estímulo más respuestas). Página 49.

Gráfico 2.2. Curvas de calibración para cada condición experimental (ES: estímulo solo, ER: estímulo más respuestas, RD: recuerdo. RN: reconocimiento). Página 51.

Gráfico 3.1. Distribución de los juicios de aprendizaje para DI. Porcentaje de respuestas para cada nivel de juicio (ES: estímulo solo. ER: estímulo más respuestas). Página 59.

Gráfico 3.2. Distribución de los juicios de aprendizaje para DE. Porcentaje de respuestas para cada nivel de juicio (ES: estímulo solo. ER: estímulo más respuestas). Página 60.

Gráfico 3.3. Curvas de calibración de DI (ES: estímulo solo, ER: estímulo más respuestas, RD: recuerdo. RN: reconocimiento). Página 62.

Gráfico 3.4. Curvas de calibración de DE (ES: estímulo solo, ER: estímulo más respuestas, RD: recuerdo. RN: reconocimiento). Página 63.

BIBLIOGRAFÍA

- Algarabel, S., Ruiz, J. C. y San Martín, J. (1988). The University of Valencia's computerized word pool. *Behavior Research Methods, Instruments, and Computers*, 20, 398-403.
- Arbuckle, T. Y. y Cuddy, L. L. (1969). Discrimination of item strength at time of presentation. *Journal of Experimental Psychology*, 81, 126-131.
- Atkinson, R. C., Herrmann, D. J. y Westcourt, K. T. (1974). Search processes in recognition memory. En R. L. Solso (Eds.), *Theories in cognitive psychology: The Loyola Symposium* (págs. 101-146). Hillsdale, NJ: Erlbaum.
- Atkinson, R. C. y Juola, J. F. (1974). Search and decision processes in recognition memory. En D. H. Krantz, R. C. Atkinson, R. D. Luce y P. Suppes (Eds.), *Contemporary developments in mathematical psychology*, Vol. 1: Learning, memory and thinking. San Francisco: Freeman.
- Barrett, A. B., Dienes, Z. y Seth, A. K. (2013). Measures of Metacognition on Signal-Detection Theoretic Models. *Psychological Methods*, 18, 535-552.
- Begg, I., Duft, S., Lalonde, P. Melnick, R. y Sanvito, J. (1989). Memory predictions are based on ease of processing. *Journal of Memory and Language*, 28, 610-632.
- Begg, I. y Rowe, E. J. (1972). Continuous judgments of words frequency and familiarity. *Journal of Experimental Psychology*, 95, 48-54.
- Benjamin, A. S. y Díaz, M. (2008). Measurement of relative mnemonic accuracy. En J. Dunlosky y R.A. Bjork (Eds.), *Handbook of metamemory and memory* (págs. 73-94). Hove: Psychology Press.
- Brown, R. y McNeill, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Behavior*, 5, 325-337.
- Buchler, N. G., Light, L. L. y Reder, L. M. (2008). Memory for items and associations: Distinct representations and processes in associative recognition. *Journal of Memory and Language*, 59, 183-199.
- Bugg, J. M. y McDaniel, M. A. (2012). Selective Benefits of Question Self-Generation and Answering for Remembering Expository Text. *Journal of Educational Psychology*, 104, 922-931.
- Burke, D., Mackay, D. G., Worthley, J. S., y Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults?. *Journal of Memory and Language*, 30, 237-246.
- Castel, A. D. y Craik, F. I. M. (2003). The Effects of Aging and Divided Attention on Memory for Item and Associative Information. *Psychology and Aging*, 18, 873-885.
- Clark, S. E. (1992). Word frequency effects in associative and item recognition. *Memory & Cognition*, 20, 231-243.
- Clark, S. E. y Shiffrin, R. M. (1992). Cueing effects and associative information in recognition memory. *Memory & Cognition*, 20, 580-598.

- Craik, F. I. M. (2002). Levels of processing: Past, present... and future?. *Memory*, *10*, 305–318.
- Craik, F. I. M., y Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671–684.
- Coane, J. H., Balota, D. A., Dolan, P. O. y Jacoby, L. L. (2011). Not all sources of familiarity are created equal: the ease of word frequency and repetition in episodic recognition. *Memory & Cognition*, *39*, 791-805.
- Connor, L. T., Hertzog, C. y Dunlosky, J. (1997). Age-Related differences in absolute but not relative metamemory accuracy. *Psychology and Aging*, *12*, 50-71.
- Criss, A. H. y Shiffrin, R. M. (2004). Interactions between study task, study time, and the low-frequency hit rate advantage in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 778–786.
- Davis, R., Sutherland, N. S. y Judd, B. R. (1961). Information content in recognition and recall. *Journal of Experimental Psychology*, *61*(5), 422-429.
- Dougherty, M. R., Scheck, P., Nelson, T. O. y Narens, L. (2005). Using the past to predict the future. *Memory & Cognition*, *33*, 1096-1115.
- Dunlosky, J. y Nelson, T. O. (1992). Importance of the kind of cue for judgments of learning (JOL) and the delayed-JOL effect. *Memory & Cognition*, *20*, 374-380.
- Dunlosky, J. y Nelson, T. O. (1994). Does the sensitivity of judgments of learning (JOLs) to the effects of various study activities depend on when the JOLs occur?. *Journal of Memory and Language*, *33*, 545-565.
- Dunlosky, J. y Nelson, T. O. (1997). Similarity between the Cue for Judgments of Learning (JOL) and the Cue for Test is not the Primary Determinant of JOL Accuracy. *Journal of Memory and Language*, *36*, 34-49.
- Dunlosky, J., Rawson, K. A. y Middleton, E. L. (2005). What constrains the accuracy of metacomprehension judgments? Testing the transfer-appropriate-monitoring and accessibility hypotheses. *Journal of Memory and Language* *52*, 551–565
- Dunlosky, J. y Thiede, K. W. (2013). *Metamemory*. New York: Oxford University Press.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive developmental inquiry. *American Psychologist*, *34*, 906–911.
- García, J. L., González, M. A. y Ballesteros, B. (2001). *Introducción a la Investigación en Educación (Tomo II)*. Madrid: UNED.
- Glanzer, M y Adams, J. K. (1990). The mirror effect in recognition memory: data and theory. *Journal of Experimental Psychology*, *16*, 5-16.
- Gorman, A. M. (1961). Recognition memory for nouns as a function of abstractedness and frequency. *Journal of Experimental Psychology*, *61*, 23-39.
- Graf, P. y Ryan, L. (1990). Transfer-appropriate processing for implicit and explicit

- memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 978-992.
- Green D. M. y Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Greene, R. L. y Thapar, A. (1994). Mirror effect in frequency discrimination. *Journal of Experimental Psychology*, 20, 946-952.
- Gregg, V. (1976). Word frequency, recognition and recall. En J. Brown (Eds.), *Recall and recognition*, (págs. 183-216), London: Wiley.
- Hall, J. (1954). Learning as a function of word-frequency. *American Journal of Psychology*, 67, 138-140.
- Hart, J. T. (1965). Memory and the feeling of knowing experience. *Journal of Educational Psychology*, 56, 208-216.
- Hart, J. T. (1967). Memory and the memory-monitoring process. *Journal of Verbal Learning and Verbal Behavior*, 6, 685-691.
- Haskins, A. L., Yonelinas, A. P., Quamme, J. R. y Ranganath, C. (2008). Perirhinal Cortex Supports Encoding and Familiarity-Based Recognition of Novel Associations. *Neuron*, 59, 554-560.
- Hertwig, R., Herzog, S. M., Schooler, L. J. y Reimer, T. (2008). Fluency Heuristic: A Model of How the Mind Exploits a By-Product of Information Retrieval. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 34, 1191-1206.
- Hines, J. C., Tournon, D. R. y Hertzog, C. (2009). Metacognitive Influences on Study Time Allocation in an Associative Recognition Task: An Analysis of Adult Age Differences. *Psychology and Aging*, 24, 462-475.
- Howell, D. C. (2013). *Statistical Methods for Psychology*. Ontario: Cengage Learning.
- Jacoby, L. L. y Dallas, M. (1981). On the relationship of between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, 110, 306-340.
- James, L. E. y Burke, D. M. (2000). Phonological priming effects on word retrieval and tip of the tongue experiences in young and older adults. *Journal of Experimental Psychology: Learning Memory and Cognition*, 26, 1378-1391.
- Jang, Y., Pashler, H. y Huber, D. E. (2014). Manipulations of choice familiarity in multiple-choice testing support a retrieval practice account of the testing effect. *Journal of Educational Psychology*: 106, 435-447.
- Kahneman, D. (2012). *Pensar rápido, pensar despacio* (Traducción de Chamorro, J.). Barcelona: Debate (Obra original publicada en 2011).
- Kahneman, D. y Frederick, S. (2005). *A Model of Heuristic Judgment*. Cambridge handbook of thinking and reasoning. En Holyoak, Keith J. (Eds.); Morrison, Robert G. (Eds.) págs.267-293. New York, NY, US: Cambridge University Press.
- Kahneman, D. y Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237-251.

- Karpicke, J. D., Lehman, M. y Aue, W. R. (2014). Retrieval-based learning: An episodic-context account. En B.H. Ross (Eds.), *The psychology of learning and motivation 61*, 237-284. San Diego, CA: Elsevier.
- Kennedy, M. R. T. y Nawrocki, M. D. (2003). Delayed predictive accuracy of narrative recall after traumatic brain injury: Salience and explicitness. *Journal of Speech, Language, and Hearing Research*, 46, 98–112.
- Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review*, 100, 609-639.
- Koriat, A. (1994). Memory's knowledge of its own knowledge. The accessibility account of the feeling of knowing. En J. Metcalfe y A.P. Shimamura (Eds.), *Metacognition. Knowing about knowing* (115-135) Cambridge, MA: MIT Press.
- Koriat, A. (1997). Monitoring one's own knowledge during study: a cue utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, 126, 349-370.
- Koriat, A. (2012). The relationships between monitoring, regulation and performance. *Learning and Instruction* 22, 296-298.
- Koriat, A. y Levy-Sadot, R. (2001). The Combined Contributions of the Cue-Familiarity and Accessibility Heuristics to Feelings of Knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 34-53.
- Koriat, A. y Ma'ayan, H. (2005). The effects of encoding fluency and retrieval fluency on judgments of learning. *Journal of Memory and Language* 52, 478–492.
- Koriat, A. y Nussinson, R. (2009). Attributing study effort to data-driven and goal-driven effects: Implications for metacognitive judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 338–1343
- Koriat, A., Nussinson, R. y Ackerman, R. (2014). Judgments of Learning Depend on How Learners Interpret Study Effort. *Journal of Experimental Psychology: Learning, Memory and Cognition* 40(6), 1624–1637.
- Lee, H., Fell, J. y Axmacher, N. (2013). Electrical engram: how deep brain stimulation affects memory. *Trends in Cognitive Sciences*, 17, 574-584.
- Leonesio, R. J. y Nelson, T. O. (1990). Do different metamemory judgments tap the same underlying aspects of memory? *Journal of Experimental Psychology: Learning, Memory and Cognition*, 16, 464-470.
- Li, S.-C., Naveh-Benjamin, M., y Lindenberger, U. (2005). Aging neuromodulation impairs associative binding. *Psychological Science*, 16, 445–450.
- Lockhart, R. S. (2002). Levels of processing, transfer-appropriate processing, and the concept of robust encoding. *Memory*, 10, 397-403.
- Lohnas L. J., Polyn S. M. y Kahana M. J. (2015). Expanding the Scope of Memory Search: Modeling Intralist and Interlist Effects in Free Recall. *Psychological Review*, 122, 337-363.
- Lovelace, E. A. (1984). Metamemory: Monitoring future recallability during study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10,

756-766.

- Luce, R. D. (1963). A threshold theory for simple detection experiments. *Psychological Review*, 70, 61-79.
- Macmillan, N. A. y Creelman, C. D. (1991). *Detection Theory: A User's Guide*. Nueva York: Cambridge University Press.
- Malmberg, K. J. (2008). Recognition Memory: A Review of the Critical Findings and an Integrated Theory for Relating Them. *Cognitive Psychology*, 57, 335-384.
- Malmberg, K. J., Holden, J. E. y Shiffrin, R. M. (2004). Modeling the effects of repetitions, similarity, and normative word frequency on old–new recognition and judgments of frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 319-331.
- Malmberg, K. J. y Nelson, T. O. (2003). The word frequency effect for recognition memory and the elevated-attention hypothesis. *Memory & Cognition*, 31, 35-43.
- Malmberg, K. J. y Xu, J. (2007). On the flexibility and fallibility of associative memory. *Memory & Cognition*, 35, 545-556.
- Malmberg, K. J., Zeelenberg, R. y Shiffrin, R. M. (2004). Turning up the noise or turning down the volume? On the nature of the impairment of episodic recognition memory by midazolam. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 540-549.
- Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, 87, 252-271.
- Mandler, G. (1991). Your face looks familiar but I can't remember your name: A review of dual process theory. En W. E. Hockley and S. Lewandowsky (Eds.), *Relating theory and data: Essays on human memory in honor of B.B. Murdock* (págs. 227–248). Hillsdale, NJ: Erlbaum.
- Maniscalco, B. y Lau, H. (2012). A signal-detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition*, 21, 422-430.
- Masson, M. E. J. y Rotello, C. R. (2009). Sources of bias in the Goodman-Kruskal gamma coefficient measure of association: Implications for studies of metacognitive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 509-527.
- McBride, D. M. y Abney, D. H. (2012). A Comparison of Transfer-Appropriate Processing and Multi-Process Frameworks for Prospective Memory Performance. *Experimental Psychology*, 59, 190–198.
- McDaniel, M. A., Kowitz, M. D., y Dunay, P. K. (1989). Altering memory through recall: The effects of cue-guided retrieval processing. *Memory & Cognition*, 17, 423-434.
- Metcalf, J. (1993). Novelty monitoring, metacognition, and control in a composite holographic associative recall model: Implications for Korsakoff amnesia. *Psychological Review*, 100, 3-22.

- Metcalfe, J. y Finn, B. (2008). Familiarity and Retrieval Processes in Delayed Judgments of Learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 1084–1097.
- Metcalfe J., Schwartz, B. L. y Joaquim, S. G. (1993). The cuefamiliarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 851–864.
- Metcalfe, J. y Shimamura, A. P. (1994). *Metacognition: Knowing about knowing*. Cambridge, MA: MIT Press.
- Morris, C. D., Bransford, J. D., y Franks, J. J. (1977). Levels of processing versus transfer-appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, 16, 519–533.
- Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1170–1187
- Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, 95, 109-133
- Nelson, T. O. y Dunlosky, J. (1991). When people's judgments of learnig (JOLs) are extremely accurate at predicting subsequent recall: The "Delayed-JOL effect". *Psychological Science*, 2, 267-270.
- Nelson, T. O., Gerler, D. y Narens, L. (1984). Accuracy of feeling of knowing judgments for predicting perceptual identification and relearning. *Journal of Experimental Psychology: General*, 113, 282-300.
- Nelson, T. O. y Narens, L. (1990). Metamemory: A theoretical framework and new findings. *The Psychology of Learning and Motivation*, 26, 125-173.
- Nelson, T. O., y Narens, L. (1994). Why investigate metacognition? En J. Metcalfe & A.P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (págs. 1–25). Cambridge, MA: MIT Press.
- Nelson, T. O., Narens, L., y Dunlosky, J. (2004). A revised methodology for research on metamemory: Pre-judgment recall and monitoring (PRAM). *Psychological Methods*, 9, 53-69.
- Nomi, J., Rhodes, M. y Cleary, A. (2013). Emotional facial expressions differentially influence predictions and performance for face recognition. *Cognition and Emotion*, 27, 141-149.
- Old, S. R., y Naveh-Benjamin, M. (2008). Differential effects of aging on item and associative measures of memory: A meta-analysis. *Psychology and Aging*, 23, 104–118.
- Park, H. y Rugg, M. D. (2008). The relationship between study processing and the effects of cue congruency at retrieval: MRI support for transfer appropriate processing. *Cerebral Cortex*, 18, 868-875.
- Parks, C. M. (2013). Transfer-Appropriate Processing in Recognition Memory: Perceptual and Conceptual Effects on Recognition Memory Depend on Task Demands. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 4, 1280–1286.

- Parks, C.M. y Yonelinas, A.P (2015). The importance of unitization for familiarity-based learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41, 881-903.
- Perfect, T. J. y Schwartz, B. L. (2002). *Applied metacognition*. Cambridge: Cambridge University Press.
- Poirier, M., Nairne, J. S., Morin, C. y Zimmermann, F. G. S. (2012). Memory as Discrimination: Challenge to the Encoding–Retrieval Match Principle. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 16–29
- Raghunathan, T. E., Rosenthal, R. y Rubin, D. B. (1996). Comparing correlated but nonoverlapping correlations. *Psychological Methods*, 1, 178-183.
- Ratcliff, R., Thapar, A. y McKoon, G. (2011). Effects of Aging and IQ on Item and Associative Memory. *Journal of Experimental Psychology: General*, 140, 464–487.
- Reder, L. M. y Ritter, F. E. (1992). What determines initial feeling of knowing? Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 435-452.
- Rhodes, M. G. y Tauber, S. K. (2011). The Influence of Delaying Judgments of Learning on Metacognitive Accuracy: A Meta-Analytic Review. *Psychological Bulletin*, 137, 131–148.
- Roediger, H. L. y Guynn, M. J. (1996). Retrieval processes. En E. L. Bjork & R. A. Bjork (Eds.), *Human memory* (págs. 197-236). San Diego: Academic Press.
- Rose, N. S. y Craik, F. I. M. (2012). A Processing Approach to the Working Memory/Long-Term Memory Distinction: Evidence From the Levels-of-Processing Span Task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 1019-1029.
- Rose, N. S., Myerson, J., Roediger, H. L. y Hale, S. (2010). Similarities and differences between working memory and long-term memory: Evidence from the levels-of-processing span task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36, 471–483.
- Ruiz, M. (2003). *Procedimientos y simulaciones en psicología de la memoria*. Madrid: UNED.
- Ruiz, M. (2004). *Las caras de la memoria*. Madrid: Pearson-Prentice Hall.
- Ruiz, M. y Arroyo, C. (2014). *CorCor.ods: A Calc workbook for testing the difference between two non-independent overlapping or non-overlapping correlations*. Manuscrito no publicado.
- Ruiz, M. y Arroyo, C. (2015). *Measuring judgment of learning resolution for predicting recognition through TDS-dependent indices*. Manuscrito en preparación.
- Ruiz, M. y Arroyo, C. (2016). JOLer: A Java standalone application for simulating the Weaver & Kelemen's judgments of learning (JOL) model. *Anales de Psicología*, 32(3), págs. por determinar.
- Ruiz-Vargas, J. M. y Cuevas, I (1999). Priming perceptivo versus priming

- conceptual y efectos de los niveles de procesamiento sobre la memoria implícita. *Psicothema*, *11*, 853-871
- Salas, C. R., Minakata, K. y Kelemen, W. L. (2011). Walking before study enhances free recall but not judgement-of-learning magnitude. *Journal of Cognitive Psychology*, *23*, 507-513.
- Sauvage, M. M., Fortin, N. J., Owens, C. B., Yonelinas, A. P. y Eichenbaum, H. (2008). Recognition memory: opposite effects of hippocampal damage on recollection and familiarity. *Nature Neuroscience*, *11*, 16-18.
- Schacter, D. L. (1983). Feeling of knowing in episodic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*, 39-54.
- Schulman, A. I. (1976). Memory for rare words previously rated for familiarity. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 301-307.
- Schwartz, B. L. (1994). Sources of information in metamemory: judgments of learning and feelings of knowing. *Psychonomic Bulletin and Review*, *3*, 357-375.
- Schwartz, B. L. y Metcalfe, J. (1992). Cue familiarity but not target retrievability enhances feeling-of-knowing judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 1074-1083.
- Soderstrom, N. C., Clark, C. T., Halamish, V. y Bjork, E. L. (2015). Judgments of learning as memory modifiers. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 553-559.
- Son, L. K. y Metcalfe, J. (2005). Judgments of learning: Evidence for atwo-stage model. *Memory & Cognition*, *33*, 1116-1129.
- Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological bulletin*, *87*, 245-251.
- Sumby, W. H. (1963). Word frequency and serial position effects. *Journal of Verbal Learning and Verbal Behavior*, *1*, 443-450.
- Susser, J. A., Mulligan, N. W y Besken, M. (2013). The effects of list composition and perceptual fluency on judgments of learning (JOLs). *Memory & Cognition*, *41*, 1000-1011.
- Thiede, K. W. y Dunlosky, J. (1994). Delaying studeants' metacognitive monitoring improves their accuracy in predicting their recognition performance. *Journal of Educational Psychology*, *86*, 290-302.
- Thomson, D. M. y Tulving, E. (1970). Associative encoding and retrieval: weak and strong cues. *Journal of Experimental Psychology*, *86*, 255-262.
- Tulving, E. (1985). Memory and Consciousness. *Canadian Psychologist*, *5*, 1-13.
- Tulving, E. y Thomson, D. M. (1971). Retrieval processes in recognition memory: effects of associative context. *Journal of Experimental Psychology*, *87*, 116-124.
- Tulving, E. y Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*, 352-373.

- Tversky, A. y Kahneman, D. (1974). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207-232.
- Undorf, M. y Erdfelder, E. (2011). Judgments of Learning Reflect Encoding Fluency: Conclusive Evidence for the Ease-of-Processing Hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 1264–1269.
- Van Overschelde, J. y Nelson, T. O. (2006). Delayed judgments of learning cause both a decrease in absolute accuracy (calibration) and an increase in relative accuracy (resolution). *Memory & Cognition*, 34, 1527–1538.
- Verde, M. F. (2004). Associative interference in recognition memory: A dual-process account. *Memory & Cognition*, 32, 1273-1283.
- Weaver III, C. A. y Kelemen, W. L. (2003). Processing similarity does not improve metamemory: evidence against transfer-appropriate monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 1058-1065.
- Weinfurt, K. P. (2000). Repeated-measures analyses: ANOVA, MANOVA, and HLM. En L.G. Grimm & P.R. Yarnold (Eds.), *Reading and Understanding more Multivariate Statistics*, (págs. 317-361). Washington: American Psychological Association.
- Whittlesea, B. W. A. (2003). On the construction of behavior and subjective experience: The production and evaluation of performance. En J. Bowers & C. Marsolek (Eds.), *Rethinking implicit memory* (págs. 239-260). Oxford: Oxford University Press.
- Xu, J. y Malmberg, K. J. (2007). Modeling the effects of verbal and nonverbal pair strength on associative recognition. *Memory & Cognition*, 35, 526-544.
- Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 1341–1354.
- Yonelinas, A. P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. *Memory & Cognition*, 25, 747-763.
- Yonelinas, A. P. (2007). Remembering: Defining and measuring. En Roediger HL III, Dudai Y, & Fitzpatrick SM (Eds.), *Science of Memory: Concepts* (págs. 233-236). New York: Oxford University Press.
- Yonelinas, A. P., Aly, M., Wang, W. y Koen, J. D. (2010). Recollection and familiarity: Examining controversial assumptions and new directions. *Hippocampus*, 20, 1178–1194.
- Zechmeister, E.B. y Niberg, S.E. (1982). *Human memory and introduction to research and theory*. Monterey, CA: Brooks/Cole.

APÉNDICE 1.

Análisis estadístico del Experimento 1

1.1: ANOVA de medidas repetidas de cada participante del rendimiento en las pruebas.

Pruebas de efectos intra-sujetos.

Medida: Rendimiento_Pruebas
Greenhouse-Geisser

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación	Eta al cuadrado parcial	Parámetro de no centralidad
Bloque	14856,702	2,740	5422,116	13,691	,000	,186	37,513
Bloque * Sesión	6904,411	5,480	1259,920	3,181	,007	,096	17,434
Error(Bloque)	65109,283	164,401	396,039				
Juicios	7715,848	1,000	7715,848	21,048	,000	,260	21,048
Juicios * Sesión	427,099	2,000	213,550	,583	,562	,019	1,165
Error(Juicios)	21994,749	60,000	366,579				
Prueba	232015,117	1,000	232015,117	226,160	,000	,790	226,160
Prueba * Sesión	1262,044	2,000	631,022	,615	,544	,020	1,230
Error(Prueba)	61553,373	60,000	1025,890				
Bloque * Juicios	552,701	2,899	190,660	,653	,577	,011	1,894
Bloque * Juicios * Sesión	5532,098	5,798	954,180	3,270	,005	,098	18,957
Error(Bloque*Juicios)	50757,262	173,933	291,821				
Bloque * Prueba	2144,977	2,852	752,188	2,144	,100	,035	6,114
Bloque * Prueba * Sesión	8112,231	5,703	1422,375	4,055	,001	,119	23,124
Error(Bloque*Prueba)	60023,660	171,099	350,812				
Juicios * Prueba	1578,720	1,000	1578,720	3,404	,070	,054	3,404
Juicios * Prueba * Sesión	1082,372	2,000	541,186	1,167	,318	,037	2,334
Error(Juicios*Prueba)	27823,801	60,000	463,730				
Bloque * Juicios * Prueba	1305,276	2,801	466,068	1,404	,245	,023	3,933
Bloque * Juicios * Prueba * Sesión	7602,463	5,601	1357,285	4,089	,001	,120	22,905
Error(Bloque*Juicios*Prueba)	55773,393	168,037	331,912				

a Calculado con alfa = ,05

Pruebas de los efectos inter-sujetos

Medida: Rendimiento_Pruebas

Variable transformada: Promedio

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación	Eta al cuadrado parcial	Parámetro de no centralidad
Intersección	765744,477	1	765744,477	123,313	,000	,673	123,313
Sesión	4189,233	2	2094,616	,337	,715	,011	,675
Error	372587,044	60	6209,784				

a. Calculado con alfa = ,05

1.2: ANOVA de medidas repetidas de cada participante de la magnitud de los juicios.

Pruebas de efectos intra-sujetos.

Medida: Magnitud_Juicios
Greenhouse-Geisser

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación	Eta al cuadrado parcial	Parámetro de no centralidad
Bloque	19777,450	2,713	7291,141	26,449	,000	,306	71,745
Bloque * Sesión	923,059	5,425	170,147	,617	,700	,020	3,348
Error(Bloque)	44864,889	162,752	275,664				
Juicios	121195,250	1,000	121195,250	136,701	,000	,695	136,701
Juicios * Sesión	798,167	2,000	399,084	,450	,640	,015	,900
Error(Juicios)	53194,280	60,000	886,571				
Prueba	59,968	1,000	59,968	,487	,488	,008	,487
Prueba * Sesión	72,064	2,000	36,032	,293	,747	,010	,585
Error(Prueba)	7386,023	60,000	123,100				
Bloque * Juicios	5143,187	2,519	2041,588	9,798	,000	,140	24,684
Bloque * Juicios * Sesión	1240,652	5,038	246,238	1,182	,321	,038	5,954
Error(Bloque*Juicios)	31494,454	151,153	208,362				
Bloque * Prueba	2261,636	2,963	763,325	6,076	,001	,092	18,002
Bloque * Prueba * Sesión	7451,295	5,926	1257,443	10,009	,000	,250	59,311
Error(Bloque*Prueba)	22333,682	177,772	125,631				
Juicios * Prueba	60,463	1,000	60,463	,474	,494	,008	,474
Juicios * Prueba * Sesión	118,052	2,000	59,026	,463	,632	,015	,926
Error(Juicios*Prueba)	7647,204	60,000	127,453				
Bloque * Juicios * Prueba	784,447	2,859	274,336	2,331	,079	,037	6,665
Bloque * Juicios * Prueba * Sesión	1766,646	5,719	308,915	2,625	,020	,080	15,010
Error(Bloque*Juicios*Prueba)	20193,440	171,566	117,701				

a Calculado con alfa = ,05

Pruebas de los efectos inter-sujetos

Medida: Magnitud_Juicios

Variable transformada: Promedio

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación	Eta al cuadrado parcial	Parámetro de no centralidad
Intersección	1648361,730	1	1648361,730	442,989	,000	,881	442,989
Sesión	33291,758	2	16645,879	4,473	,015	,130	8,947
Error	223260,049	60	3721,001				

a. Calculado con alfa = ,05

1.3: ANOVA de medidas repetidas de las correlaciones gamma de cada participante.

Pruebas de efectos intra-sujetos.

Medida: Correlación_Gamma
Greenhouse-Geisser

Fuente	Suma de cuadrados tipo III	Gl	Media cuadrática	F	Significación	Eta al cuadrado parcial	Parámetro de no centralidad
Juicios	,164	1,000	,164	2,250	,143	,062	2,250
Juicios * Sesión	,034	2,000	,017	,235	,792	,014	,470
Error(Juicios)	2,486	34,000	,073				
Prueba	4,819	1,000	4,819	86,393	,000	,718	86,393
Prueba * Sesión	,393	2,000	,197	3,525	,041	,172	7,049
Error(Prueba)	1,896	34,000	,056				
Juicios * Prueba	,199	1,000	,199	4,391	,044	,114	4,391
Juicios * Prueba * Sesión	,134	2,000	,067	1,476	,243	,080	2,951
Error(Juicios* Prueba)	1,539	34,000	,045				

a Calculado con alfa = ,05

Pruebas de los efectos inter-sujetos

Medida: Correlación_Gamma
Variable transformada: Promedio

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación	Eta al cuadrado parcial	Parámetro de no centralidad
Intersección	50,547	1	50,547	335,151	,000	,908	335,151
Sesión	,212	2	,106	,703	,502	,040	1,406
Error	5,128	34	,151				

a Calculado con alfa = ,05

1.4: Análisis estadísticos de las rectas de regresión.

Recta ESRD

Resumen del modelo: ESRD

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,898(a)	,806	,782	7,88392

a Variables predictoras: (Constante), JUICIOS

ANOVA(b)

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	2068,139	1	2068,139	33,273	,000(a)
	Residual	497,249	8	62,156		
	Total	2565,388	9			

a Variables predictoras: (Constante), JUICIOS

b Variable dependiente: ESRD

Coefficientes(c)

Modelo		Coefficients no estandarizados		Coefficientes estandarizados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	-9,328	5,005		-1,864	,099
	JUICIOS	,501	,087	,898	5,768	,000

c Variable dependiente: ESRD

Recta ESRN

Resumen del modelo: ESRN

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,900(a)	,810	,786	9,03662

a Variables predictoras: (Constante), JUICIOS

ANOVA(b)

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	2782,497	1	2782,497	34,074	,000(a)
	Residual	653,285	8	81,661		
	Total	3435,781	9			

a Variables predictoras: (Constante), JUICIOS

b Variable dependiente: ESRN

Coefficientes(c)

Modelo		Coefficients no estandarizados		Coefficientes estandarizados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	25,628	5,737		4,467	,002
	JUICIOS	,581	,099	,900	5,837	,000

c Variable dependiente: ESRN

Recta ERRD

Resumen del modelo: ERRD

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,963(a)	,928	,919	4,28349

a Variables predictoras: (Constante), JUICIOS

ANOVA(b)

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	1898,300	1	1898,300	103,459	,000(a)
	Residual	146,786	8	18,348		
	Total	2045,086	9			

a Variables predictoras: (Constante), JUICIOS

b Variable dependiente: ERRD

Coeficientes(c)

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	-7,716	2,719		-2,837	,022
	JUICIOS	,480	,047	,963	10,171	,000

c Variable dependiente: ERRD

Recta ERRN

Resumen del modelo: ERRN

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,965(a)	,930	,922	5,63154

a Variables predictoras: (Constante), JUICIOS

ANOVA(b)

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	3394,577	1	3394,577	107,036	,000(a)
	Residual	253,714	8	31,714		
	Total	3648,292	9			

a Variables predictoras: (Constante), JUICIOS

b Variable dependiente: ERRN

Coeficientes(c)

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	11,243	3,575		3,145	,014
	JUICIOS	,641	,062	,965	10,346	,000

c Variable dependiente: ERRN

Para comparar las rectas de regresión de nuestro experimento debemos tener en cuenta que las puntuaciones de las diferentes rectas proceden de los mismos individuos. Se trata, por tanto, de puntuaciones diferentes -no solapadas- pero correlacionadas, ya que proceden de la misma muestra. Pues bien, Raghunathan y otros (1996; véase también Steiger, 1980) recomiendan calcular la puntuación ZPF cuando necesitamos una prueba de la significación estadística de la diferencia entre dos correlaciones relacionadas no solapadas. Esta prueba no está aún incorporada a los paquetes estadísticos usuales (SPSS o SAS), por lo que hemos desarrollado unas macros en Calc, con las que analizamos los datos de esta tesis (Ruiz y Arroyo, 2014).

A continuación se muestran las puntuaciones ZPF de las diferencias de las rectas de regresión tomadas de dos en dos y su significación:

	ZPF (N=10)	p(ZPF)
ESRD-ERRD	1.40	.081
ESRD-ESRN	0.03	.489
ESRN-ERRN	1.34	.090
ERRD-ERRN	0.05	.481

APÉNDICE 2.

Análisis estadístico del Experimento 2

2.1: ANOVA de medidas repetidas de cada participante del rendimiento en las pruebas.

Pruebas de efectos intra-sujetos.

Medida: Rendimiento_Pruebas
Greenhouse-Geisser

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación	Eta al cuadrado parcial	Parámetro de no centralidad
Bloque	2572,422	1,000	2572,422	5,479	,023	,097	5,479
Bloque * Sesión	1335,525	1,000	1335,525	2,845	,098	,053	2,845
Error(Bloque)	23943,196	51,000	469,474				
Juicios	8758,492	1,000	8758,492	32,431	,000	,389	32,431
Juicios * Sesión	186,165	1,000	186,165	,689	,410	,013	,689
Error(Juicios)	13773,164	51,000	270,062				
Pruebas	35869,407	1,000	35869,407	113,788	,000	,691	113,788
Pruebas * Sesión	20,350	1,000	20,350	,065	,800	,001	,065
Error(Pruebas)	16076,715	51,000	315,230				
Bloque * Juicios	1187,618	1,000	1187,618	7,166	,010	,123	7,166
Bloque * Juicios * Sesión	113,404	1,000	113,404	,684	,412	,013	,684
Error(Bloque*Juicios)	8452,319	51,000	165,732				
Bloque * Pruebas	1519,851	1,000	1519,851	7,559	,008	,129	7,559
Bloque * Pruebas * Sesión	247,733	1,000	247,733	1,232	,272	,024	1,232
Error(Bloque*Pruebas)	10254,887	51,000	201,076				
Juicios * Pruebas	1090,305	1,000	1090,305	6,093	,017	,107	6,093
Juicios * Pruebas * Sesión	1348,167	1,000	1348,167	7,534	,008	,129	7,534
Error(Juicios*Pruebas)	9125,628	51,000	178,934				
Bloque * Juicios * Pruebas	44,299	1,000	44,299	,198	,659	,004	,198
Bloque * Juicios * Pruebas * Sesión	560,022	1,000	560,022	2,497	,120	,047	2,497
Error(Bloque*Juicios*Pruebas)	11438,406	51,000	224,282				

a Calculado con alfa = ,05

Pruebas de los efectos inter-sujetos

Medida: Rendimiento_Pruebas
Variable transformada: Promedio

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación	Eta al cuadrado parcial	Parámetro de no centralidad
Intersección	655383,150	1	655383,150	165,212	,000	,764	165,212
Sesión	12,081	1	12,081	,003	,956	,000	,003
Error	202312,657	51	3966,915				

a. Calculado con alfa = ,05

2.2: ANOVA de medidas repetidas de cada participante de la magnitud de los juicios.

Pruebas de efectos intra-sujetos.

Medida: Magnitud_Juicios
Greenhouse-Geisser

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación	Eta al cuadrado parcial	Parámetro de no centralidad
Bloque	72,371	1,000	72,371	,283	,597	,006	,283
Bloque * Sesión	14,150	1,000	14,150	,055	,815	,001	,055
Error(Bloque)	13052,864	51,000	255,939				
Juicios	51998,080	1,000	51998,080	120,496	,000	,703	120,496
Juicios * Sesión	128,157	1,000	128,157	,297	,588	,006	,297
Error(Juicios)	22008,246	51,000	431,534				
Pruebas	24959,602	1,000	24959,602	125,602	,000	,711	125,602
Pruebas * Sesión	72,226	1,000	72,226	,363	,549	,007	,363
Error(Pruebas)	10134,712	51,000	198,720				
Bloque * Juicios	35,168	1,000	35,168	,235	,630	,005	,235
Bloque * Juicios * Sesión	157,292	1,000	157,292	1,049	,310	,020	1,049
Error(Bloque*Juicios)	7644,391	51,000	149,890				
Bloque * Pruebas	232,081	1,000	232,081	2,962	,091	,055	2,962
Bloque * Pruebas * Sesión	1027,266	1,000	1027,266	13,111	,001	,205	13,111
Error(Bloque*Pruebas)	3995,826	51,000	78,350				
Juicios * Pruebas	15,481	1,000	15,481	,157	,694	,003	,157
Juicios * Pruebas * Sesión	237,304	1,000	237,304	2,404	,127	,045	2,404
Error(Juicios*Pruebas)	5035,355	51,000	98,732				
Bloque * Juicios * Pruebas	58,443	1,000	58,443	,814	,371	,016	,814
Bloque * Juicios * Pruebas * Sesión	,343	1,000	,343	,005	,945	,000	,005
Error(Bloque*Juicios*Pruebas)	3659,566	51,000	71,756				

a Calculado con alfa = ,05

Pruebas de los efectos inter-sujetos

Medida: Magnitud_Juicios

Variable transformada: Promedio

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación	Eta al cuadrado parcial	Parámetro de no centralidad
Intersección	986931,364	1	986931,364	431,902	,000	,894	431,902
Sesión	1694,639	1	1694,639	,742	,393	,014	,742
Error	116539,114	51	2285,081				

a. Calculado con alfa = ,05

2.3: ANOVA de medidas repetidas de las correlaciones gamma de cada participante.

Pruebas de efectos intra-sujetos.

Medida: Correlación_Gamma
Greenhouse-Geisser

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación	Eta al cuadrado parcial	Parámetro de no centralidad
Juicios	,775	1,000	,775	8,414	,006	,185	8,414
Juicios * Sesión	,057	1,000	,057	,623	,435	,017	,623
Error(Juicios)	3,406	37,000	,092				
Pruebas	3,298	1,000	3,298	31,530	,000	,460	31,530
Pruebas * Sesión	,005	1,000	,005	,046	,831	,001	,046
Error(Pruebas)	3,870	37,000	,105				
Juicios * Pruebas	,976	1,000	,976	20,502	,000	,357	20,502
Juicios * Pruebas * Sesión	,085	1,000	,085	1,782	,190	,046	1,782
Error(Juicios*Pruebas)	1,761	37,000	,048				

a Calculado con alfa = ,05

Pruebas de los efectos inter-sujetos

Medida: Correlación_Gamma
Variable transformada: Promedio

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación	Eta al cuadrado parcial	Parámetro de no centralidad
Intersección	61,773	1	61,773	622,144	,000	,944	622,144
Sesión	,056	1	,056	,567	,456	,015	,567
Error	3,674	37	,099				

a Calculado con alfa = ,05

2.4: Análisis estadísticos de las rectas de regresión.

Recta ESRD

Resumen del modelo: ESRD

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,912(a)	,831	,810	11,71387

a Variables predictoras: (Constante), JUICIOS

ANOVA(b)

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	5396,218	1	5396,218	39,327	,000(a)
	Residual	1097,719	8	137,215		
	Total	6493,937	9			

a Variables predictoras: (Constante), JUICIOS

b Variable dependiente: ESRD

Coefficientes(c)

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	-13,664	7,437		-1,837	,103
	JUICIOS	,809	,129	,912	6,271	,000

c Variable dependiente: ESRD

Recta ESRN

Resumen del modelo: ESRN

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,943(a)	,890	,876	7,42557

a Variables predictoras: (Constante), JUICIOS

ANOVA(b)

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	3552,451	1	3552,451	64,427	,000(a)
	Residual	441,113	8	55,139		
	Total	3993,563	9			

a Variables predictoras: (Constante), JUICIOS

b Variable dependiente: ESRN

Coefficientes(c)

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	22,639	4,714		4,802	,001
	JUICIOS	,656	,082	,943	8,027	,000

c Variable dependiente: ESRN

Recta ERRD

Resumen del modelo: ERRD

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,948(a)	,898	,885	6,44718

a Variables predictoras: (Constante), JUICIOS

ANOVA(b)

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	2924,151	1	2924,151	70,349	,000(a)
	Residual	332,529	8	41,566		
	Total	3256,681	9			

a Variables predictoras: (Constante), JUICIOS

b Variable dependiente: ERRD

Coeficientes(c)

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	-8,288	4,093		-2,025	,077
	JUICIOS	,595	,071	,948	8,387	,000

c Variable dependiente: ERRD

Recta ERRN

Resumen del modelo: ERRN

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,960(a)	,921	,911	5,95080

a Variables predictoras: (Constante), JUICIOS

ANOVA(b)

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	3306,365	1	3306,365	93,369	,000(a)
	Residual	283,296	8	35,412		
	Total	3589,661	9			

a Variables predictoras: (Constante), JUICIOS

b Variable dependiente: ERRN

Coeficientes(c)

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	24,849	3,778		6,577	,000
	JUICIOS	,633	,066	,960	9,663	,000

c Variable dependiente: ERRN

A continuación se muestran las puntuaciones ZPF (véase Apéndice 1, sección 1.4) de las diferencias de las rectas de regresión tomadas de dos en dos y su significación:

	ZPF (N=10)	p(ZPF)
ESRD-ERRD	0.99	.162
ESRD-ESRN	0.89	.187
ESRN-ERRN	0.44	.328
ERRD-ERRN	0.35	.364

2.5: Ecuación de Spearman-Brown que pone en relación la fiabilidad y el número de elementos (García y otros, 2001).

$$R_{xx} = \frac{n \cdot r_{xx}}{1 + (n-1) r_{xx}}$$

Donde:

- n: cociente entre el número de elementos que va a tener la prueba y el número de elementos inicialmente establecidos.
- r_{xx} : Coeficiente de fiabilidad inicial.
- R_{xx} : Coeficiente de fiabilidad final.

APÉNDICE 3.

Análisis estadístico del Experimento 3

3.1: ANOVA de medidas repetidas de cada participante del rendimiento en las pruebas.

Pruebas de efectos intra-sujetos.

Medida: Rendimiento_Pruebas
Greenhouse-Geisser

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación	Eta al cuadrado parcial	Parámetro de no centralidad
Bloque	5750,028	1,000	5750,028	15,815	,000	,198	15,815
Bloque * tipo de distractor	210,887	1,000	210,887	,580	,449	,009	,580
Error(Bloque)	23269,585	64,000	363,587				
Juicios	6102,087	1,000	6102,087	23,243	,000	,266	23,243
Juicios * tipo de distractor	171,532	1,000	171,532	,653	,422	,010	,653
Error(Juicios)	16802,122	64,000	262,533				
Pruebas	45155,699	1,000	45155,699	103,012	,000	,617	103,012
Pruebas * tipo de distractor	5520,262	1,000	5520,262	12,593	,001	,164	12,593
Error(Pruebas)	28054,654	64,000	438,354				
Bloque * Juicios	144,152	1,000	144,152	,733	,395	,011	,733
Bloque * Juicios * tipo de distractor	179,926	1,000	179,926	,915	,342	,014	,915
Error(Bloque*Juicios)	12588,171	64,000	196,690				
Bloque * Pruebas	225,964	1,000	225,964	1,101	,298	,017	1,101
Bloque * Pruebas * tipo de distractor	63,927	1,000	63,927	,312	,579	,005	,312
Error(Bloque*Pruebas)	13131,023	64,000	205,172				
Juicios * Pruebas	753,602	1,000	753,602	3,578	,063	,053	3,578
Juicios * Pruebas * tipo de distractor	163,787	1,000	163,787	,778	,381	,012	,778
Error(Juicios*Pruebas)	13479,899	64,000	210,623				
Bloque * Juicios * Pruebas	477,757	1,000	477,757	3,112	,082	,046	3,112
Bloque * Juicios * Pruebas * tipo de distractor	36,091	1,000	36,091	,235	,629	,004	,235
Error(Bloque*Juicios*Pruebas)	9824,936	64,000	153,515				

a Calculado con alfa = ,05

Pruebas de los efectos inter-sujetos

Medida: Rendimiento_Pruebas
Variable transformada: Promedio

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación	Eta al cuadrado parcial	Parámetro de no centralidad
Intersección	1293076,055	1	1293076,055	307,559	,000	,828	307,559
tipo de distractor	30358,462	1	30358,462	7,221	,009	,101	7,221
Error	269076,386	64	4204,319				

a Calculado con alfa = ,05

3.2: ANOVA de medidas repetidas de cada participante de la magnitud de los juicios.

Pruebas de efectos intra-sujetos.

Medida: Magnitud_Juicios
Greenhouse-Geisser

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación	Eta al cuadrado parcial	Parámetro de no centralidad
Bloque	739,524	1,000	739,524	4,777	,033	,069	4,777
Bloque * tipo de distractor	97,026	1,000	97,026	,627	,431	,010	,627
Error(Bloque)	9908,624	64,000	154,822				
Juicios	77193,459	1,000	77193,459	214,036	,000	,770	214,036
Juicios * tipo de distractor	1419,230	1,000	1419,230	3,935	,052	,058	3,935
Error(Juicios)	23081,959	64,000	360,656				
Pruebas	42273,184	1,000	42273,184	274,854	,000	,811	274,854
Pruebas * tipo de distractor	623,935	1,000	623,935	4,057	,048	,060	4,057
Error(Pruebas)	9843,353	64,000	153,802				
Bloque * Juicios	62,964	1,000	62,964	,684	,411	,011	,684
Bloque * Juicios * tipo de distractor	,549	1,000	,549	,006	,939	,000	,006
Error(Bloque*Juicios)	5888,654	64,000	92,010				
Bloque * Pruebas	26,133	1,000	26,133	,230	,633	,004	,230
Bloque * Pruebas * tipo de distractor	1,596	1,000	1,596	,014	,906	,000	,014
Error(Bloque*Pruebas)	7261,817	64,000	113,466				
Juicios * Pruebas	162,016	1,000	162,016	1,385	,244	,021	1,385
Juicios * Pruebas * tipo de distractor	36,595	1,000	36,595	,313	,578	,005	,313
Error(Juicios*Pruebas)	7486,240	64,000	116,973				
Bloque * Juicios * Pruebas	31,015	1,000	31,015	,449	,505	,007	,449
Bloque * Juicios * Pruebas * tipo de distractor	8,024	1,000	8,024	,116	,734	,002	,116
Error(Bloque*Juicios*Pruebas)	4419,726	64,000	69,058				

a Calculado con alfa = ,05

Pruebas de los efectos inter-sujetos

Medida: Magnitud_Juicios
Variable transformada: Promedio

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación	Eta al cuadrado parcial	Parámetro de no centralidad
Intersección	1608578,413	1	1608578,413	816,847	,000	,927	816,847
tipo de distractor	164,931	1	164,931	,084	,773	,001	,084
Error	126032,190	64	1969,253				

a Calculado con alfa = ,05

3.3.1: ANOVA mixto de tres factores (2 x (2x2xS)) de medidas repetidas de las correlaciones gamma de cada participante.

Pruebas de efectos intra-sujetos.

Medida: Correlación_Gamma
Greenhouse-Geisser

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación	Eta al cuadrado parcial
Juicios	,024	1,000	,024	,183	,670	,003
Juicios * tipo de distractor	,388	1,000	,388	2,942	,092	,051
Error(Juicios)	7,255	55,000	,132			
Pruebas	5,584	1,000	5,584	31,152	,000	,362
Pruebas * tipo de distractor	,005	1,000	,005	,027	,869	,000
Error(Pruebas)	9,858	55,000	,179			
Juicios * Pruebas	,136	1,000	,136	1,151	,288	,020
Juicios * Pruebas * tipo de distractor	,601	1,000	,601	5,077	,028	,085
Error(Juicios*Pruebas)	6,514	55,000	,118			

a Calculado con alfa = ,05

Pruebas de los efectos inter-sujetos

Medida: Correlación_Gamma
Variable transformada: Promedio

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación	Eta al cuadrado parcial
Intersección	86,739	1	86,739	528,752	,000	,906
tipo de distractor	,001	1	,001	,004	,948	,000
Error	9,022	55	,164			

a Calculado con alfa = ,05

3.3.2: ANOVA de medidas repetidas de las correlaciones gamma de cada participante separado por *tipo de distractor*.

Distractores Internos (DI). Pruebas de efectos intra-sujetos.

Medida: Correlación_Gamma
Greenhouse-Geisser

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación	Eta al cuadrado parcial	Parámetro de no centralidad
Juicios	,320	1,000	,320	3,326	,079	,103	3,326
Error(Juicios)	2,788	29,000	,096				
Pruebas	3,124	1,000	3,124	20,437	,000	,413	20,437
Error(Pruebas)	4,433	29,000	,153				
Juicios * Pruebas	,087	1,000	,087	,894	,352	,030	,894
Error(Juicios*Pruebas)	2,825	29,000	,097				

a Calculado con alfa = ,05

Distractores Externos (DE). Pruebas de efectos intra-sujetos.

Medida: Correlación_Gamma
Greenhouse-Geisser

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación	Eta al cuadrado parcial	Parámetro de no centralidad
Juicios	,104	1,000	,104	,604	,444	,023	,604
Error(Juicios)	4,467	26,000	,172				
Pruebas	2,497	1,000	2,497	11,969	,002	,315	11,969
Error(Pruebas)	5,425	26,000	,209				
Juicios * Pruebas	,622	1,000	,622	4,386	,046	,144	4,386
Error(Juicios*Pruebas)	3,689	26,000	,142				

a Calculado con alfa = ,05

3.4: Análisis estadísticos de las rectas de regresión.

Recta ESRD

Resumen del modelo: ESRD-DI

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,833(a)	,693	,655	13,80437

a Variables predictoras: (Constante), JUICIOS

ANOVA(b)

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	3444,231	1	3444,231	18,074	,003(a)
	Residual	1524,486	8	190,561		
	Total	4968,717	9			

a Variables predictoras: (Constante), JUICIOS

b Variable dependiente: ESRD

Coeficientes(c)

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	-9,980	8,764		-1,139	,288
	JUICIOS	,646	,152	,833	4,251	,003

c Variable dependiente: ESRD

Resumen del modelo: ESRD-DE

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,917(a)	,841	,821	10,64447

a Variables predictoras: (Constante), JUICIOS

ANOVA(b)

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	4777,122	1	4777,122	42,162	,000(a)
	Residual	906,438	8	113,305		
	Total	5683,559	9			

a Variables predictoras: (Constante), JUICIOS

b Variable dependiente: ESRD

Coeficientes(c)

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	-1,627	6,758		-,241	,816
	JUICIOS	,761	,117	,917	6,493	,000

c Variable dependiente: ESRD

Recta ESRN

Resumen del modelo: ESRN-DI

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,853(a)	,728	,694	14,51411

a Variables predictoras: (Constante), JUICIOS

ANOVA(b)

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	4510,988	1	4510,988	21,414	,002(a)
	Residual	1685,276	8	210,659		
	Total	6196,264	9			

a Variables predictoras: (Constante), JUICIOS

b Variable dependiente: ESRN

Coeficientes(c)

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	16,061	9,214		1,743	,119
	JUICIOS	,739	,160	,853	4,627	,002

c Variable dependiente: ESRN

Resumen del modelo: ESRN-DE

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,424(a)	,180	,078	10,98066

a Variables predictoras: (Constante), JUICIOS

ANOVA(b)

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	212,005	1	212,005	1,758	,221(a)
	Residual	964,600	8	120,575		
	Total	1176,604	9			

a Variables predictoras: (Constante), JUICIOS

b Variable dependiente: ESRN

Coeficientes(c)

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	68,359	6,971		9,806	,000
	JUICIOS	,160	,121	,424	1,326	,221

c Variable dependiente: ESRN

Recta ERRD

Resumen del modelo: ERRD-DI

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,924(a)	,853	,835	9,10826

a Variables predictoras: (Constante), JUICIOS

ANOVA(b)

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	3855,410	1	3855,410	46,473	,000(a)
	Residual	663,683	8	82,960		
	Total	4519,093	9			

a Variables predictoras: (Constante), JUICIOS

b Variable dependiente: ERRD

Coeficientes(c)

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	-6,928	5,782		-1,198	,265
	JUICIOS	,684	,100	,924		

c Variable dependiente: ERRD

Resumen del modelo: ERRD-DE

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,934(a)	,872	,855	9,49682

a Variables predictoras: (Constante), JUICIOS

ANOVA(b)

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	4893,947	1	4893,947	54,263	,000(a)
	Residual	721,516	8	90,190		
	Total	5615,463	9			

a Variables predictoras: (Constante), JUICIOS

b Variable dependiente: ERRD

Coeficientes(c)

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	-7,683	6,029		-1,274	,238
	JUICIOS	,770	,105	,934		

c Variable dependiente: ERRD

Recta ERRN

Resumen del modelo: ERRN-DI

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,804(a)	,646	,601	14,63921

a Variables predictoras: (Constante), JUICIOS

ANOVA(b)

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	3124,243	1	3124,243	14,578	,005(a)
	Residual	1714,452	8	214,306		
	Total	4838,695	9			

a Variables predictoras: (Constante), JUICIOS

b Variable dependiente: ERRN

Coeficientes(c)

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	13,340	9,294		1,435	,189
	JUICIOS	,615	,161	,804		

c Variable dependiente: ERRN

Resumen del modelo: ERRN-DE

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,873(a)	,763	,733	10,21325

a Variables predictoras: (Constante), JUICIOS

ANOVA(b)

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	2682,860	1	2682,860	25,720	,001(a)
	Residual	834,484	8	104,311		
	Total	3517,344	9			

a Variables predictoras: (Constante), JUICIOS

b Variable dependiente: ERRN

Coeficientes(c)

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	43,583	6,484		6,722	,000
	JUICIOS	,570	,112	,873		

c Variable dependiente: ERRN

A continuación se muestran las puntuaciones ZPF (véase Apéndice 1, sección 1.4) de las diferencias de las rectas de regresión tomadas de dos en dos y su significación para las dos condiciones:

	DI		DE	
	ZPF (N=10)	p(ZPF)	ZPF (N=10)	p(ZPF)
ESRD-ERRD	0,98	.163	0.33	.370
ESRD-ESRN	0.21	.418	2.40	.008
ESRN-ERRN	0.41	.341	2.16	.015
ERRD-ERRN	1.16	.122	0.81	.209