

III Congreso Internacional de Investigación y Práctica Profesional en Psicología XVIII Jornadas de Investigación Séptimo Encuentro de Investigadores en Psicología del MERCOSUR. Facultad de Psicología - Universidad de Buenos Aires, Buenos Aires, 2011.

Evaluación de tareas de razonamiento silogístico en alumnos de psicología.

Attorresi, Horacio Félix, Galibert, María Silvia, Blum, G. Diego y Abal, Facundo Juan Pablo.

Cita:

Attorresi, Horacio Félix, Galibert, María Silvia, Blum, G. Diego y Abal, Facundo Juan Pablo (2011). *Evaluación de tareas de razonamiento silogístico en alumnos de psicología. III Congreso Internacional de Investigación y Práctica Profesional en Psicología XVIII Jornadas de Investigación Séptimo Encuentro de Investigadores en Psicología del MERCOSUR. Facultad de Psicología - Universidad de Buenos Aires, Buenos Aires.*

Dirección estable: <https://www.aacademica.org/000-052/910>

ARK: <https://n2t.net/ark:/13683/eRwr/SFT>

Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto. Acta Académica fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: <https://www.aacademica.org>.

EVALUACIÓN DE TAREAS DE RAZONAMIENTO SILOGÍSTICO EN ALUMNOS DE PSICOLOGÍA

Attorresi, Horacio Félix; Galibert, María Silvia; Blum, G. Diego; Abal, Facundo Juan Pablo
UBACyT, Facultad de Psicología, Universidad de Buenos Aires

RESUMEN

Se presentan los resultados de la administración de 31 ítems de silogismos a 403 alumnos de la Facultad de Psicología, UBA. Cada ítem tiene cinco alternativas semi-abiertas, sólo una correcta. Las distribuciones de frecuencias para cada alternativa revelaron el tipo de dificultad, especialmente para tres ítems que resultaron extremadamente difíciles y poco correlacionados. Se ajustó el modelo logístico de tres parámetros a 25 ítems, lo que constituyó una prueba de confiabilidad 0,77, tendiente a difícil con discriminación moderada. La cantidad media de respuestas correctas a dicha prueba fue 9,6 (DS=3,8). La interacción entre edad y sexo resultó significativa al 10% ($p=0,0742$) con una diferencia a favor de los varones entre los jóvenes, con un tamaño de efecto moderado de 0,49, y a favor de las mujeres entre los mayores con tamaño de efecto de 0,68. Estudios del Funcionamiento Diferencial del Ítem (DIF) con el método Mantel-Haenszel revelaron uno solo con DIF pronunciado a favor de las mujeres. En cuanto a los niveles de dificultad de los ítems según modo y figura, los principales resultados fueron convergentes con los reportados por la bibliografía, como por ejemplo el efecto del sesgo de la figura y el efecto atmósfera.

Palabras clave

Silogismo Modelo logístico de tres parámetros Funcionamiento diferencial del ítem Sesgo de la figura

ABSTRACT

EVALUATION OF SYLLOGISTIC REASONING TASKS AMONG STUDENTS OF PSYCHOLOGY

This study presents the results of the administration of 31 syllogisms to 403 students of Psychology of UBA. Each item has five semi-open alternatives, only one correct. Frequency distributions for each alternative showed the kind of difficulty, especially for three items that were extremely difficult and poorly correlated. The three parameter logistic model was fitted on a subset of 25 items, which made up a test with reliability 0.77; this test tends to be difficult with moderate discrimination. The mean for correct answers was 9.6 (SD = 3.8). The interaction between age and sex was significant at 10% ($p = 0.0742$) with a difference in favor of males among the young, with a moderate effect size of 0.49, and for women aged older with an effect size of 0.68. Studies of Differential Item Functioning (DIF) with the Mantel-Haenszel method revealed only one with DIF in favor of women. As for the difficulty levels of the items by mode and figure, the main results were convergent with those reported in the

literature, such as the effect of the figural bias and the atmospheric effect.

Key words

Syllogism 3Parameters logistic model Differential item functioning Figural bias

Introducción

El Razonamiento de Predicados o Silogístico estudia la comprensión de los cuantificadores universal y existencial del cálculo de predicados, los cuales en el lenguaje natural corresponden a “todos” y “algunos”. Investigadores de la psicología del razonamiento, como Santamaría (1995) y Espino, O. (2004), señalan la importancia del silogismo como representante de los estudios sobre el razonamiento deductivo y las ventajas que presenta en las tareas experimentales por las cuales se ha constituido en paradigma de la investigación en esta área. De allí la proliferación de modelos para explicar el modo en que los sujetos resuelven los silogismos; algunas teorías haciendo énfasis en explicar la génesis de los errores y otras posteriores, en un intento más abarcativo, en develar el procedimiento por el que las personas arriban a una conclusión. Entre las teorías que hacen hincapié en explicar el error en el razonamiento silogístico las hay que postulan una actitud no del todo racional a la hora de resolver las tareas y otras que preservan mejor la competencia lógica de los sujetos. Como exponente clásico del primer grupo está la teoría del “efecto atmósfera” de Woodworth y Sells (1935) y Sells (1936) que postula que los sujetos tienden a producir una conclusión según la atmósfera generada por las premisas, en cuanto a modo se refiere; esto es universal afirmativa o negativa (A o E), o particular afirmativa o negativa (I u O). Según esta teoría, si una premisa es universal y otra particular, los sujetos tenderán a formular una conclusión de modo particular, y si una es negativa, también prevalecerá la negativa en la conclusión. A partir de allí predicen el modo de la conclusión a partir del de las premisas. Por el contrario, las teorías del segundo grupo, explican el error por una interpretación equívoca de las premisas. Entre ellas se halla la de Chapman y Chapman (1959) que hipotetizan que los sujetos extraen conclusiones que serían válidas si no fuera por la conversión ilícita de las premisas universal afirmativa y particular negativa; es decir, el error no está en el modo en que se extrae la conclusión desde las premisas, sino en una interpretación errónea de las mismas

por conversión ilícita; por ejemplo pueden considerar equivalente “Todos los A son B” con “Todos los B son A” o “Algunos A no son B” con “Algunos B no son A”. En esta misma línea se hallan Ceraso y Provitera (1971) y Revlin y Leirer (1978). A partir de mediados de los 70, comenzaron a surgir modelos para explicar el proceso de deducción en sí mismo y no necesariamente desde el error. Erickson y Traub desarrollan modelos basados sobre la interpretación de las premisas postulando que las personas utilizan los diagramas de Euler en la resolución de los silogismos (Erickson, 1974, 1978; Traub y Erickson, 1975). Otros modelos se basan sobre el efecto de la figura; entre ellos los de Dickstein (1975, 1976, 1978, 1980, 1981) y Johnson-Laird y Steedman (1978). En Johnson-Laird y Byrne (1991) se desarrolla la representación en términos de los modelos mentales de todos los conectivos lógicos. En Johnson-Laird (1983) y Johnson-Laird y Bara (1984) se desarrolla la teoría de los modelos mentales para los silogismos donde se determinan cuatro etapas secuenciales: 1) interpretación inicial de las premisas, 2) combinación de las interpretaciones en un solo modelo, 3) formación de una conclusión informativa y 4) búsqueda de modelos alternativos que puedan refutar la conclusión.

Las investigaciones mencionadas se enmarcaron en un enfoque cognitivo. En este trabajo se aborda el tema también desde un enfoque psicométrico, proponiendo una prueba para la medición del razonamiento deductivo en su aspecto silogístico. Se examinan las respuestas a las opciones para verificar las predicciones de las principales teorías cognitivas y se compara el desempeño según el sexo con los correspondientes análisis del potencial funcionamiento diferencial de los ítems.

Método

Participantes

La muestra estuvo integrada por 403 alumnos de segundo año de la carrera de Psicología de la UBA. La prueba se administró en una clase teórica de hora y media en cada horario en que se dicta la materia. La distribución según el sexo es de 86 % de mujeres y 14% de varones. La edad promedio es de 21 con una desviación estándar de 5.

Materiales

Se administró una prueba de 31 silogismos. Como todo silogismo, consiste de la premisa mayor, la premisa menor y luego una conclusión. En esta prueba la conclusión debe obtenerse a partir de cinco opciones. Para minimizar el efecto del azar propio de la elección múltiple pero al mismo tiempo evitar las respuestas inclasificables, siguiendo a Santamaría (1989) y de manera similar a Espino et al. (2000), se optó por un formato de respuesta cerrada a saber: todos los ... son ..., algunos ... son ..., algunos ... no son ..., ningún ... es ..., no hay conclusión válida. Los sujetos deberán completar las líneas de puntos con los términos extremos de las premisas para construir la conclusión correcta si existiere o elegir la última opción. En cuanto al contenido, siguiendo la idea de Jon-

son-Laird (1978) se eligió contenido concreto para que el razonamiento se aproxime al modo en se opera en la realidad, pero neutro, tratando de minimizar las relaciones semánticas entre los términos de las premisas, para evitar sesgar las respuestas. Los factores que se han tenido en cuenta para asignar las características al material estimular son el modo y la figura. Para cada una de las cuatro figuras se consideraron los 16 modos que surgen de combinar el carácter universal / particular - afirmativo / negativo de las premisas mayor y menor. De ese modo quedan determinados, con relación a las premisas, 64 silogismos, de los cuales solo 27 tienen conclusión válida. En virtud de la reversibilidad de la proposiciones particulares afirmativas (I) y universales negativas (E) y de la permutación de premisas, en realidad 30 pares de estos 64 silogismos son lógicamente equivalentes, por lo que sólo 34 son esencialmente diferentes en cuanto a la estructura. Más precisamente, cada silogismo de la figura I tiene un equivalente en la figura IV, cuatro silogismos de la figura II tienen un equivalente en la III y dentro de la figura II y de la figura III hay a su vez cinco pares de silogismos equivalentes. Como era necesario seleccionar un subconjunto de silogismos para evitar que la prueba resultara excesivamente larga, se priorizó en la selección que estos 34 silogismos esencialmente diferentes quedaran representados. De ellos se tomaron 2 como ejemplo y se dejó uno de lado por resultar demasiado similar al del ejemplo. De modo que la prueba quedó conformada por 31 ítems: 7 de la figura I y 8 para cada figura II, III y IV. Se administraron cuatro versiones permutando los ítems de modo de compensar el posible efecto del orden.

Análisis

Para el análisis psicométrico de la prueba como tal se utilizó el programa BILOG-MG. Se ajustó el modelo logístico de 3 parámetros luego de haberse excluido los ítems que en la fase 1 presentaban correlaciones casi nulas. También se dejaron de lado los ítems que perjudicaban el ajuste global según el valor p del test χ^2 . El subconjunto de ítems a los cuales ajusta el modelo da lugar a la conformación de la prueba definitiva. A partir de la prueba resultante del ajuste se compararon los sujetos según sexo y edad categorizada, a través de un análisis de la varianza y se obtuvieron las distribuciones de frecuencia para los diversos grupos. Se estudió el funcionamiento diferencial del ítem (Differential Item Functioning, DIF) entre sexos con la prueba de Mantel-Haenszel mediante el programa de Waller (1998). Para determinar los factores de dificultad de los distintos tipos de silogismos y la consistencia de los resultados con las predicciones de las principales teorías se examinaron las distribuciones de frecuencias de las respuestas y se compararon los índices de dificultad clásico (bc) y de la TRI (b) de manera descriptiva o con los correspondiente tests de diferencia de proporciones o medias.

Resultados

Influencia de la figura, modo y validez

El análisis de la dificultad asociada a estas fuentes se

detalla sólo para los ítems que quedaron excluidos del ajuste precisamente porque su gran dificultad incidieron en su baja correlación con el resto.

En Tabla 1 se muestran los índices clásicos de dificultad *bc* y de correlación *r* según la forma de los ítems 3, 5 y 19. Se especifica a qué figura pertenecen y el modo. El tercer carácter del modo es un guión si no hay conclusión válida o en caso contrario el modo de la conclusión.

Tabla 1. Índices Característicos para los Ítems 3, 5 y 19.

Ítem	3	5	19
Modo	AO-	AEO	AI-
Figura	IV	I	IV
<i>bc</i>	.102	.017	.060
<i>r</i>	-.004	.067	.086

Para el ítem 3, la mayoritaria producción (64.7%) de la conclusión de tipo *Algunos no* por sobre la respuesta correcta *No hay conclusión válida* (10.2%) está en la dirección de lo predicho por la teoría del “efecto atmósfera”, según la cual entre una proposición universal y una particular prevalece la elección de la particular y entre una afirmativa y una negativa, la negativa (Sells, 1936; Woodworth y Sells, 1935). Se verifica el sesgo de figura (Johnson-Laird y Steedman, 1978) por el que la figura IV tiende a producir conclusiones en la dirección S - P (dirección hacia adelante) y la figura I, P - S (dirección hacia atrás), donde S representa el sujeto de la primera premisa y P el predicado de la segunda. El índice de dificultad está en un orden similar al 8% hallado por Espino y Santamaría (2000) para el modo AO en la figura IV. Estos autores consideran también el efecto del número de modelos, por el que este tipo de silogismos es clasificado como de múltiples modelos. En el ítem 5 la opción preferida *Ninguno es* (58.5%) también está entre lo predicho por la teoría del efecto atmósfera y el sesgo de la figura. La dificultad extrema coincide con los resultados de Espino y Santamaría que obtuvieron sólo un 4% de respuesta correcta para el modo AE en la figura I, clasificado como correspondiente a múltiples modelos. El ítem 19 no tiene conclusión válida y la opción más elegida fue *Algunos son* (27,4%). Una vez más los resultados corresponden a los predichos por la teoría del efecto atmósfera como también con los hallazgos de Espino y Santamaría (2000) que obtienen un 8% de respuesta correcta para el modo AI en la figura IV. Si bien se verifica el sesgo de la figura, no tan pronunciadamente como en otros casos. La razón de esto es que más de 20% eligió la reversa. Con respecto a los ítems 1, 7 y 25 que se dejaron de lado en una segunda etapa para mejorar el ajuste, la diferencia de proporciones resultó significativa para el ítem 1 a favor de la figura IV. Su estimación puntual es de .14 con un intervalo del 95% de confianza de .065 a .212. Es coincidente con Espino y Santamaría (2000) el hecho de que el modo AA del ítem 1 resulte mucho más fácil que lo ítems anteriormente analizados, tal vez por corresponder a lo que ellos clasifican como de un solo modelo. La diferencia con sus resultados está en que ellos registraron una

diferencia mucho mayor entre las figuras IV y I, más difícil para esta última, mientras que en nuestro caso dicha diferencia no es tan pronunciada. Más precisamente. Espino y Santamaría obtienen un 60% de respuestas correctas para el modo AA en la figura IV y sólo un 28% para el mismo modo en la figura I; en nuestro caso esos porcentajes están respectivamente en el orden del 70% y 56% respectivamente. Si bien este ítem resultó de dificultad media podría haberse esperado que fuera mucho más fácil en virtud tanto del sesgo de la figura como del efecto atmósfera que, en este caso, acompañan a la respuesta correcta facilitándola. Un examen de las distribuciones de frecuencias para las opciones muestra que un 24% respondió a la reversa de la conclusión, lo que va de acuerdo con los modelos que explican el error por interpretación de las premisas (Ceraso y Provitera, 1971). En cuanto al modo AE en la figura IV, presentes en el ítem 25, Espino y Santamaría (2000) registran 88% de respuesta correcta para el modo AE en la figura IV, al que clasifican como de un solo modelo. A pesar de la diferencia con nuestro 74%, en ambos casos el ítem ha resultado fácil, posiblemente favorecido por el efecto atmósfera que aquí también beneficia a la opción correcta y por la reversibilidad de la conclusión. El ítem 7 resultó difícil probablemente por el efecto de la figura que obtuvo casi un 47% repartidos de manera más o menos equitativa entre “algunos son” y “algunos no son”, por lo que el efecto atmósfera acá se verificó sólo parcialmente.

Ajuste del Modelo Logístico de 3 parámetros

Los ítems 3, 5 y 19 exhibieron correlaciones muy bajas o casi nulas y fueron excluidos del ajuste para favorecer la convergencia. Estos ítems resultaron extremadamente difíciles (Tabla 1). A fin de lograr un mejor ajuste global, también quedaron excluidos los ítems 1, 7 y 25. Sobre los ítems restantes se ajustó el modelo logístico de tres parámetros ($p = .2181$) con un criterio de convergencia de .01 habiendo sido el mayor cambio .009622. El valor del estadístico c^2 para el test global fue de 226.7 ($gl = 211$) con valor $p = .2181$. En Tabla 2 se presentan las medias con sus respectivas desviaciones estándar para las estimaciones de los parámetros del modelo.

Tabla 2. Media y Desviación Estándar de los Parámetros Estimados

PARÁMETROS	MEDIA	DESV. EST.
PSEUDOAZAR (<i>c</i>)	.115	.044
DISCRIMINACIÓN (<i>a</i>)	.670	.333
DIFICULTAD (<i>b</i>)	1.3	1.76

Baker (1985) propone una clasificación para los valores del parámetro de discriminación (*a*) que sirve de orientación para su valoración cualitativa. En Tabla 3 se muestra esta clasificación con la distribución de frecuencias.

Tabla 3. Distribución del Parámetro a

Valoración cualitativa	Nula	Muy Baja	Baja	Moderada	Alta	Muy Alta
Valores numéricos	0	.01 - .34	.35 - .64	.65 - 1.34	1.35 - 1.69	> 1.70
Frecuencia	0	3	10	11	0	1

En Tabla 4 se muestran los puntos de máxima información, la información media de los ítems y la confiabilidad computada como el cociente entre las varianzas de los valores estimados del rasgo latente y de los puntajes observados. En tabla 5 se presentan los resúmenes estadísticos para el puntaje.

Tabla 4. Valores de Información y Confiabilidad.

Punto de Máxima Información		Información Media	Confiabilidad
Abscisa	Ordenada		
1	5.845	3.247	0.765

Tabla 5. Resúmenes Estadísticos del Puntaje.

n	Media	Desv. Estándar	Coef. Variación	Mínimo	1er Cuartil	Mediana	3er Cuartil	Máximo	Asimetría	Curtosis
403	9.5533	3.7561	39.317	1	7	9	12	22	.4889	.0573

Tomando en cuenta la distribución de frecuencias de la edad, ésta se categorizó en tres grupos, a saber: de 18 a 20 años, entre 21 y 26 años y mayores de 26 años. Según un ANOVA para los puntajes resultó significativa al 10% ($p = .0742$) la interacción entre edad y sexo con una diferencia a favor de los varones entre los jóvenes, con un tamaño de efecto moderado (.49) y a favor de las mujeres entre los mayores con alto tamaño de efecto (.68). En Tabla 6 se presentan las medias y errores estándar del puntaje por edad y sexo. El ANOVA en términos del rasgo latente η^2 dieron resultados similares. Resultó significativa la edad al 6% ($p = .0515$).

Tabla 6. Medias y Errores Estándar del Puntaje por Edad y Sexo.

EDAD	Mujeres			Varones		
	Media	ES	n	Media	ES	n
18-20	9.677	.2282	266	11.500	.7936	22
21-26	8.742	.4582	66	9.824	.9028	17
27-54	9.048	.8123	21	6.500	1.5196	6

Estudios del DIF

Al estudiar el DIF sólo uno fue clasificado como Tipo C (DIF pronunciado) según el criterio del Educational Testing Service, con una medida MH D-DIF de 2.311 (ES = 0.619), a favor de las mujeres. Otros dos ítems fueron clasificados en el grupo B (sospechosos de DIF); el pri-

mero a favor de los varones con MH D-DIF de -1.962 (ES 0.676) y el segundo de las mujeres con MH D-DIF de 1.996 (ES 0.582). En el resto de los ítems no se detectó DIF.

Conclusiones

A partir de la modelización psicométrica de la prueba de silogismos ésta quedó constituida por 25 ítems. Los requerimientos psicométricos de correlación y ajuste al modelo hicieron necesario dejar de lado 6 de los silogismos, 5 de los cuales corresponden a los modos con ambas premisas universales dentro de las figuras I o IV y por ello estas combinaciones particulares no han quedado representadas en la prueba, aunque sí los correspondientes modos y figuras en otras combinaciones. Cuatro de estos ítems resultaron sumamente difíciles y dos medianamente fáciles. A pesar de estas exclusiones el espectro de situaciones posibles ha quedado suficientemente representado. Con respecto a la potencia discriminadora de los ítems y, por tanto, de la información que proporcionan habría sido deseable observar valores mayores que los hallados. Generalmente en el proceso de construcción de ítems existe la posibilidad de ir mejorando sus características psicométricas introduciéndole modificaciones en sucesivas pruebas piloto. En el caso de los silogismos esto no parece ser muy provechoso. En efecto, cada silogismo tiene una estructura rígida que lo define y la única posibilidad de variación está en el contenido; puesto que se ha elegido un contenido neutro de términos más bien desvinculados con el fin de no sesgar la conclusión, no parece tener lugar una expectativa substancial de mejoramiento. En otras palabras, aquellos que han resultado extremadamente difíciles, con correlaciones casi nulas y fueron excluidos del ajuste, tanto como los que quedaron ajustados pero con bajo parámetro a posiblemente repitan su funcionamiento a pesar de una variación del contenido porque dicho funcionamiento es debido a la estructura misma del silogismo.

Los puntajes resultaron con una acentuada asimetría positiva, lo que da cuenta de su tendencia a ser una prueba difícil. Coherentemente con esto, no hubo individuos que obtuvieran el mayor puntaje y aquellos que tienen más de veinte puntos son señalados como datos atípicos.

Al comparar las distribuciones según sexo y edad se observó una tendencia a mayores puntajes para los varones más jóvenes. Al estudiar el DIF entre sexos para ver si su existencia podía explicar la tendencia a puntuar por encima a los varones, se observó que, por el contrario, en dos de los tres ítems con DIF estaban favorecidas las mujeres. No hay una explicación evidente desde el contenido de los ítems en cuestión para dar cuenta de este comportamiento diferencial y cualquier interpretación del mismo parecería aventurada en esta instancia. Por otra parte, la muestra de varones es pequeña considerando que los métodos de detección del DIF requieren considerar las frecuencias condicionales a los distintos niveles del puntaje total, por lo que con-

vendría recoger más evidencias en futuras investigaciones antes de excluir estos ítems.

En la bibliografía referida a resultados experimentales sobre razonamiento silogístico se halla una gran variabilidad en la dificultad dependiendo del tipo de silogismo, dificultad medida en términos de los porcentajes de respuestas correctas. Por ejemplo, Johnson-Laird y Steedman (1978) reportan índices entre 5% y 100%. En este estudio, si bien también se evidenció una importante variabilidad no se hallaron ítems extremadamente fáciles, tanto en términos del índice clásico de dificultad como del parámetro *b* de la TRI pero sí extremadamente difíciles. De hecho, los ítems de esta prueba han resultado en general bastante difíciles, como lo muestran las medias de sus índices de dificultad.

En cuanto a las predicciones de Dickstein (1975, 1980) acerca de la influencia de cada figura, los resultados han sido acordes con las referidas a las figuras I y IV. En efecto, en la mayoría de los ítems, las frecuencias tendieron a concentrarse más en las opciones formuladas según el sesgo de la figura aunque el porcentaje de respuestas en dirección al sesgo no alcanza los niveles del orden del 70% hallados por Johnson Laird y Steedman (1978) sino del 55% para la figura I y 59% para la figura IV. En lo que se refiere a la influencia del modo, según lo predicho por la teoría del efecto atmósfera de Woodworth y Sells (1935) y Sells (1936), se ha verificado que en la mayoría de los ítems las opciones más elegidas son las que corresponden a los modos predichos por esta teoría. Finalmente, en cuanto a la dificultad asociada a algunas interacciones particulares de figura y modo, se hallaron resultados similares a los reportados por Santamaría (2000).

BIBLIOGRAFÍA

- Ceraso, J. y Provitera, A. (1971). "Sources of error in syllogistic reasoning". *Cognitive Psychology*, 2, 400-410. Traducción Castellana: "Fuentes de error en el razonamiento silogístico". En J. A. Delval (Comp.), *Investigaciones sobre lógica y psicología*. Madrid, 1977, Alianza Editorial.
- Chapman, L. y Chapman, J. (1959). Atmosphere effect re-examined. *Journal of Experimental Psychology*, 58, 220-226. Traducción Castellana: Nuevo examen del efecto atmósfera. En J. A. Delval (Comp.), *Investigaciones sobre lógica y psicología*. Madrid, 1977. Alianza Editorial.
- Dickstein, L. (1975). Effects of instructions and premise order on errors in syllogistic reasoning. *Journal of Experimental Psychology: Human Learning and Memory*, 104, 376-384.
- Dickstein, L. (1976). Differential difficulty of categorical syllogisms. *Bulletin of Psychonomic Society*, 8, 330-332.
- Dickstein, L. (1978). The effect of figure on syllogistic reasoning. *Memory and Cognition*, 6, 76-83.
- Dickstein, L. (1980). Inference errors in deductive reasoning. *Bulletin of the Psychonomic Society*, 6, 414-416.
- Dickstein, L. (1981). Conversion and possibility in syllogistic reasoning. *Bulletin of the Psychonomic Society*, 18, 229-232.
- Erickson, J. (1974). A set análisis theory of behaviour in formal syllogistic reasoning tasks. En Saldo, R., *Theories of Cognitive Psychology*, Erlbaum: New Jersey.
- Erickson, J. (1978). Research on syllogistic reasoning. En Revlin, R. y Mayer, R. (Comps.). *Human Reasoning*. Wiley: New York.
- Espino, O. (2004). *Pensamiento y Razonamiento*. Madrid. Pirámide.
- Espino, O. y Santamaría, C. (2000). Influencia del número de modelos y de la figura en tareas de razonamiento silogístico. *Revista de Psicología General y Aplicada* 53 (4), 701-716.
- Johnson-Laird, P. y Steedman, M. (1978). The psychology of syllogisms. *Cognitive Psychology* 10, 64 - 99.
- Johnson-Laird, P. (1983). *Mental models*. Cambridge. MA: Harvard University Press.
- Johnson-Laird, P. y Bara, B. (1984). Sillogistic inference. *Cognition*, 16, 1-61.
- Johnson-Laird, P. y Byrne, R. (1991). *Deduction*. London: LEA.
- Revlin, R. y Leirer, V. (1978). The effect of personal biases on syllogistic reasoning: racional decisions from personalizad representation. En Revlin, R. y Mayer, R. (Comps.) *Human Reasoning*, Washington, D.C.: Winston.
- Santamaría, C. (1995). *Introducción al razonamiento humano*. Madrid. Alianza Editorial.
- Sells, S. (1936). The atmosphere effect: an experimental study of reasoning. *Archives of Psychology*, 29, 3-72.
- Traub, B. y Erickson, J (1975). Determinants of difficulty in judging the validity of syllogistic arguments. *Midwestern Psychological Association*. Chicago.
- Woodworth, R y Sells, S. (1935). An atmosphere effect in syllogistic reasoning. *Journal of Experimental Psychology*, 18, 451-460.