

Estrategias y conflictos semióticos de estudiantes de psicología al evaluar un problema de independencia

Cañadas de la Fuente, Gustavo R. y Gea Serrano, M. Magdalena

Universidad de Granada

En este trabajo se analizan las estrategias utilizadas en la evaluación de la independencia en una tabla de contingencia en una muestra de 414 estudiantes de psicología en tres universidades españolas. Las estrategias se analizan desde el punto de su nivel de complejidad, según la clasificación de los niveles de dificultad propuestas por Pérez Echeverría (1990). En cada nivel de esta clasificación se estudia la corrección de las estrategias, considerándose tres tipos: (a) estrategias correctas, (b) estrategias parcialmente correctas, entendiéndose estas como las que proporciona soluciones válidas en algunos problemas, pero en otros no, y (c) estrategias incorrectas. El análisis de las estrategias incorrectas y parcialmente correctas condujo a la identificación de los diversos conflictos semióticos, los cuales vienen definido en el marco teórico del enfoque onto-semiótico (Godino, Batanero y Font, 2007; Font y Godino, 2006) que es el que utilizamos en nuestra investigación. Estos conflictos semióticos implicados en el razonamiento del estudiante fueron clasificados y comentados.

Aunque dados los datos se trata de un caso de independencia perfecta, la mayoría de los estudiantes dieron un coeficiente de asociación de tamaño moderado y un juicio positivo de la asociación. Pocas estrategias llegaron al nivel más alto de complejidad y una serie de conflictos semióticos fueron identificados.

Palabras clave: Tablas de contingencia, dificultades, juicios de asociación y estrategias

1. Introducción

A pesar de que las tablas de contingencia son comunes para presentar la información estadística y de que el juicio asociación es un tema prioritario de aprendizaje en los cursos de estadística (Zieffler, 2006; Arteaga et al., 2011), se presta poca atención a su enseñanza, en el supuesto de que su interpretación es fácil.

En este trabajo se describe parte de una investigación más amplia que tuvo como objetivo evaluar la comprensión de los estudiantes de asociación en tablas de contingencia antes de enseñar y comparar sus conocimientos con los adquiridos después de una enseñanza diseñada (Cañadas, 2012).

En concreto se analizan sólo las estrategias de los estudiantes en la parte correspondiente a la independencia, porque éste era el punto donde los estudiantes eran menos precisos en la estimación del coeficiente de asociación y, además, la mayoría de los estudiantes consideró que había asociación en los datos (Batanero, Cañadas, Estepa y Arteaga, 2011; Cañadas, Batanero, Contreras y Arteaga, 2011). Las estrategias se clasifican de acuerdo con los niveles definidos por Pérez-Echeverría (1990), y su corrección. Los resultados son comparados con los descritos por Batanero, Estepa, Godino y Green (1996) y se describen los conflictos semióticos de los estudiantes que pueden explicar sus estrategias equivocadas.

2. Investigaciones previas

La investigación sobre la asociación fue iniciada por Inhelder y Piaget (1955), que describe las estrategias utilizadas en las diferentes edades cuando se juzga asociación en tareas que eran formalmente equivalente a una tabla de contingencia 2x2 (ver Tabla 1).

Tabla 1. Formato de una tabla de contingencia 2x2

	A	No A	Total
B	a	b	a+b
No B	c	d	c+d
Total	a+c	b+d	

Más tarde, los estudios psicológicos se desarrollaron con los adultos, y mostró bajo rendimiento al juzgar asociación en estas tablas. Smedslund (1963) encuentra que algunos adultos basan su juicio sólo en una celda o mediante la comparación de a con b . Allan y Jenkins (1983) mostraron la tendencia a basar la sentencia de asociación sobre la diferencia entre una celda a y d .

Aunque Allan y Jenkins (1983) encontraron que la comparación de las diagonales ($a + d$) y ($c + b$) era común en los adultos. Jenkins y Ward (1965) señaló que esta estrategia es sólo válida en las tablas que tienen filas con iguales frecuencias marginales. Una estrategia correcta para todos los tipos de tablas, según Jenkins y Ward es la comparación de las dos probabilidades condicionales, $P(B/A)$ y $P(B/no A)$. Pérez Echevarría (1990) clasifica las estrategias que se han identificado en la investigación psicológica hasta entonces en 6 niveles de rendimiento. Niveles de 0 a 3 corresponden a los estudiantes que utilizar 0 a 3 celdas para llevar a cabo el juicio de asociación. En los niveles 4 y 5, los estudiantes utilizan todas las celdas, la diferencia es que la comparación entre las celdas es de forma aditiva en el Nivel 4 y multiplicativa en el Nivel 5.

Chapman (1967, p. 151) describió una tendencia común a la que llamó "correlación ilusoria": "*el informe de los observadores de una correlación entre dos clases de acontecimientos que, en realidad, (a) no están correlacionados, (b) se correlacionan en menor medida que lo reportado, o (c) se correlacionan en la dirección opuesta a la que se informa*". Muchos investigadores han informado de esta tendencia y sugieren que las teorías previas alteran las estimaciones de la asociación (por ejemplo, Meiser y Hewstone, 2006).

Batanero, Estepa, Godino y Green (1996) analizaron el comportamiento de 213 estudiantes de 17 años y sus estrategias en los juicios de asociación y define diferentes concepciones de la asociación: (a) la concepción causal según la cual el sujeto sólo tiene en cuenta la asociación entre variables, cuando se puede explicar por la presencia de una relación causa - efecto; concepción unidireccional, por el cual el estudiante no acepta una asociación inversa, y la concepción local, donde se deduce la asociación de sólo una parte de los datos. Otros autores (por ejemplo, Contreras, 2012) encuentran errores en el cálculo de probabilidades a partir de tablas de contingencia.

Nuestra investigación tiene como objetivo ir más allá en el análisis de las estrategias en el caso de que los datos muestran independencia perfecta, ya que esta fue la tarea en la que los estudiantes se desempeñaron peor en el estudio de Batanero et al. (1996), y en nuestro estudio anterior (Batanero, Cañadas y Arteaga, 2011). También complementa el tipo de análisis efectuados en la investigación anterior con el fin de identificar los posibles conflictos semióticos de los estudiantes en la evaluación de la independencia. Este término se ha tomado de Godino, Batanero, y Font (2007). Los autores utilizan el término conflicto semiótico como: "*interpretación disparidades entre el estudiante y el significado en la institución de las matemáticas*" (p. 133).

3. Método

La muestra incluyó a 414 estudiantes en su primer año de estudios de psicología de tres universidades españolas: Almería (115), Granada (237) y Huelva (62), todos ellos tomando un curso de introducción a la estadística. La tarea encomendada a los estudiantes es una adaptación de Batanero et al. (1996), cambiando a un contexto de diagnóstico psicológico, sin variar los datos. Las muestras incluyeron todos los estudiantes matriculados en el curso, la diferencia en el tamaño de la muestra se debe al tamaño de la Universidad. Los estudiantes no habían estudiado asociación, pero si las estadísticas de secundaria.

En la parte (a), los estudiantes tienen que proporcionar un juicio asociación. Hay dos categorías de respuestas: el estudiante considera que las variables están o no relacionadas. La estimación del coeficiente de asociación se deduce en la escala numérica (0-1) en la segunda parte del ítem. Por último, un análisis cualitativo se utilizó para identificar las estrategias utilizadas por los estudiantes y sus conflictos semióticos. La clasificación de las estrategias fue realizada por dos miembros del equipo, en caso de desacuerdo, se discutió con los otros miembros del equipo hasta que se alcanzó un acuerdo.

4. Resultados y discusión

4.1. Juicios de asociación y estimación de la asociación

En la Tabla 2 se presenta el porcentaje de estudiantes que consideran la existencia de asociación y el valor medio de la estimación del coeficiente de asociación. La mayoría de los estudiantes indicaron la existencia de asociación a pesar de que los datos corresponden a un caso de independencia perfecta. Este juicio es consistente con la estimación del coeficiente de asociación, que es de tamaño moderado en promedio. Las diferencias en la estimación de la media o en el porcentaje de alumnos a juzgar asociación (en el test de chi-cuadrado) entre las tres universidades no fueron estadísticamente significativas, esto sugiere que las respuestas de los estudiantes fueron similares. Los resultados pueden explicarse por la correlación ilusoria (Chapman, 1967) ya que los datos contradice las teorías previas, así como por la concepción causal de la asociación, reportados por Batanero et al. (1996).

Tabla 2. Medias estimadas de asociación y juicios de asociación

	Almería (n=115)	Granada (n=237)	Huelva (n=62)	Total (n=414)
Media estimada	0.51	0.47	0.44	0.47
Número (y porcentaje) de estudiantes que consideran que hay asociación	87 (75.7)	194 (81.9)	42 (67.7)	323 (78.1)

4.2. Niveles de estrategias

Con el fin de explicar los resultados anteriores hemos analizado las estrategias de los estudiantes y las prácticas matemáticas que intervienen en estas estrategias. En primer lugar, las estrategias de los estudiantes fueron clasificadas de acuerdo a su corrección, en tres grupos: (a) estrategias correctas, (b) estrategias parcialmente correctas que produce respuestas correctas sólo para tablas específicas, y (c) estrategias incorrectas. Esta clasificación se cruzó con los niveles de dificultad propuestas por Pérez Echeverría (1990):

Nivel 0. El estudiante no usa la información de la tabla y sólo se tuvieron en cuenta sus propias teorías previas sobre la asociación, la correlación ilusoria (Chapman y Chapman, 1969; Murphy y Medin, 1985).

Nivel 1. Cuando el alumno sólo utiliza una celda de la tabla, por lo general es la celda cuando ambas variables están presentes y tienen un mayor impacto en la atención (Smedlund, 1963; Beyth y Maron, 1982; Shaklee y Mins, 1982).

Nivel 2. Algunos estudiantes utilizaron dos celdas, por ejemplo, en comparación a con b ó a con c , por lo que dedujeron asociación de una sola distribución condicional. Otros estudiantes comparan las celdas de frecuencia máxima con la de frecuencia mínima (Batanero et al., 1996).

Nivel 3: En esta estrategia los estudiantes utilizaron tres celdas, por ejemplo, comparar la celda a con las celdas b y c . En general, estos estudiantes descartaron la celda d (Batanero et al, 1996.). Todas las estrategias de los niveles 1 a 3 son incorrectas ya que los estudiantes sólo utilizan una parte de los datos, mostrando una concepción local de la asociación (Batanero et al., 1996).

Nivel 4: Se basa en la comparación de forma aditiva de las cuatro celdas. Un ejemplo es la comparación de la suma de las diagonales ($a + d$) con ($b + c$). Esta estrategia fue encontrada por Allan y Jenkins (1983). Esta estrategia podría proporcionar una buena solución cuando las frecuencias marginales (número de personas con y sin insomnio) son iguales, de acuerdo con Shaklee (1983). En otro ejemplo, los estudiantes compararon dos distribuciones condicionales en forma aditiva.

Nivel 5 utiliza todas las celdas con comparaciones multiplicativas, pero todavía puede ser incorrecta o parcialmente correcta. Por ejemplo, una estrategia equivocada consiste en calcular todas las frecuencias relativas conjuntas y compararlas. Un ejemplo de una estrategia parcialmente correcta está en asumir que todas las frecuencias relativas en la tabla deben ser idénticas. Por último, entre las estrategias correctas encontramos estudiantes que compararon dos distribuciones condicionales, por ejemplo, $a/(a+b)$ con $c/(c+d)$. Otra estrategia correcta es la comparación de las posibilidades a favor y en contra de B para cada valor de A; que fue descrito por Batanero et al. (1996).

Tabla 3. Frecuencias (y porcentajes) de estrategias por niveles

	Incorrecta	Parcialmente correcta	Correcta	Total
Nivel 0	13 (100.0)			13 (3.1)
Nivel 1	73 (100.0)			73 (17.6)
Nivel 2	108 (100.0)			108 (26.1)
Nivel 3	27 (100.0)			27 (6.5)
Nivel 4	27 (26.2)	76 (73.8)		103 (24.9)
Nivel 5	20 (26.3)	10 (13.2)	46 (60.5)	76 (18.4)
No response	14 (100.0)			14 (3.4)
Total	282 (68.1)	86 (20.8)	46 (11.1)	414 (100)

En la Tabla 3 se presenta la frecuencia de las respuestas en la clasificación anterior. Sólo el 11,1% de los estudiantes utilizan estrategias correctas y el 20,8% de los cuales son parcialmente correctas. Los estudiantes tienden a utilizar los niveles 2 ó 4, en los cuales ninguna de las estrategias es correcta, y además había un gran porcentaje de estudiantes que no utilizaron toda la información de las celdas, ya que sus estrategias eran de nivel 3 o inferior.

En el nivel 4, cerca de 30% de los estudiantes compara frecuencias absolutas entre ellas, una estrategia errónea descrita por Batanero et al. (1996) y aproximadamente 60% utiliza las cuatro celdas con comparaciones aditivas, una estrategia descrita por Inhelder y Piaget (1955), también se encontrada por Batanero et al. (1996) en estudiantes de secundaria. Finalmente, la mayoría de estrategias de nivel 5 eran correctas ya sea comparando los odds ratios o las distribuciones condicionales, estrategia propuesta por Jenkins y Ward (1965) y también se encuentra en Batanero et al. (1996). La precisión en la

estimación del coeficiente de asociación aumenta con la estrategia de corrección y las diferencias fueron estadísticamente significativas en la prueba de ANOVA.

4.3. Conflictos semióticos

El análisis de las estrategias incorrectas y parcialmente correctas condujo a la identificación de diversos conflictos semióticos las cuales solo mencionamos por motivos de la extensión, que se clasifican a continuación, según ya se trate de las disparidades en los estudiantes el significado asignado a la asociación o independencia (atribuyendo propiedades incorrectas a estos conceptos) u otros conflictos:

1. Propiedades incorrectas atribuidas a la asociación. Se encontraron las siguientes:
 - La identificación de la concepción causal de la asociación por Batanero et al, (1996) se encontró en las estrategias de nivel 0.
 - Suponer que la asociación se puede deducir a partir de sólo una parte de los datos (concepción local).
 - Asumir que es posible deducir de las comparaciones aditivas asociación, una estrategia que debería haber sido superado, de acuerdo con Inhelder y Piaget (1951) en la etapa de las operaciones formales.
 - Suponer que la asociación se puede deducir a partir de sólo una distribución condicional, que también fue descrito por Inhelder y Piaget (1951).
 - Pensar que la diferencia de las frecuencias absolutas condicionales es suficiente para mantener a la asociación, un error que se encontró en Smedlund (1963) y Shaklee y Mins (1982).
 - Considerar que la asociación se puede deducir de la diferencia entre las sumas de las diagonales de la tabla. Esa estrategia fue considerada como correcta por Piaget e Inhelder, pero Allan y Jenkins (1983) sugirió que no funciona en el caso general.
 - Si se asume que $a > d$ en caso de asociación. Estos estudiantes consideraron que sólo estas dos celdas influían en la asociación.
2. Propiedades incorrectas atribuidas a la asociación. La principal fue la confusión entre las ideas de independencia y equiprobabilidad.
3. Otros conflictos, como basar el juicio asociación en su propia opinión.

5. Discusión y consecuencias de la enseñanza

La mayoría de los estudiantes de psicología en nuestro estudio han juzgado asociación en una tarea donde había independencia perfecta, debido al fenómeno de correlación ilusoria, afectando a su precisión en la estimación del coeficiente de asociación.

Los resultados de nuestro estudio eran peores en comparación con Batanero et al. (1996). Estos autores no informaron acerca de la estimación de la asociación por parte de sus estudiantes, en nuestro estudio, la estimación es consistente con el criterio de asociación. Los resultados fueron muy parecidos en todas las universidades participantes. En cuanto a los conceptos descritos por estos autores, se observó la concepción causal y la local. Ya que no pudimos comprobar la estabilidad de estas concepciones, se utiliza en su lugar la idea de conflicto semiótico, que sólo implica una interpretación errónea de una expresión matemática, un concepto, propiedad o procedimiento.

Nuestra lista de conflictos semióticos es más amplia que la lista de los conceptos descritos por los autores, como nuevos conflictos relacionados con las ideas de la asociación y la independencia fueron identificados. Por ejemplo, asumir que la asociación puede ser juzgada a partir de sólo una parte de los datos, otro conflicto frecuente fue el supuesto de que la asociación puede ser juzgada a partir de las frecuencias absolutas. Puesto que los conflictos semióticos no asumen una fuerte convicción por parte de los estudiantes, es posible cambiarlos con una instrucción adecuada y la identificación de estos conflictos es un primer paso para corregir su razonamiento erróneo.

Estas razones y nuestros resultados sugieren la necesidad de una mayor investigación sobre la enseñanza de la asociación, la concepción causal y el efecto de correlación ilusoria no parece mejorar con la enseñanza tradicional (Batanero, Godino, y Estepa, 1998). Nuestro propósito es continuar este trabajo mediante el diseño de una enseñanza alternativa a las actividades que se enfrentan los estudiantes con sus conflictos y ayudarles a superarlos.

Agradecimientos: Proyecto EDU2010-14947 y Grupo FQMN-126 (Junta de Andalucía).

Referencias

- Allan, L. G. y Jenkins, H.M. (1983). The effect of representations of binary variables on judgment of influence, *Learning and Motivation*, 14, 381-405.
- Arteaga, P., Batanero, C., Cañadas, G. y Contreras, J. M. (2011). Las tablas y gráficos estadísticos como objetos culturales, *Números* 76, 55-67.
- Batanero, C., Cañadas, G., Estepa, A. y Arteaga, P. (2011) Psychology students' estimation of association. *Statistique et Enseignement*, 3 (1), 3-18.
- Batanero, C., Estepa, A., Godino, J. D. y Green, D. (1996) Intuitive strategies and preconceptions about association in contingency tables. *Journal for Research in Mathematics Education*, 27(2), 151-169.
- Batanero, C., Godino, J. D. y Estepa, A. (1998). Building the meaning of statistical association through data analysis activities. Research Forum. In A. Olivier & K. Newstead (Eds.), *Proceedings of the 22nd Conference of the International Group for the Psychology of Mathematics Education* (v.1, pp. 221-242). University of Stellenbosch.
- Beyth-Marrow, R. (1982). Perception of correlation reexamined. *Memory and cognition*, 10 (6), 511-519.
- Cañadas, G. (2012). *Comprensión intuitiva y aprendizaje formal de las tablas de contingencia en alumnos de psicología*. Tesis doctoral. Universidad de Granada.
- Cañadas, G., Batanero, C., Contreras, J. M. y Arteaga, P. (2011). Estrategias en el estudio de la asociación en tablas de contingencia por estudiantes de psicología. *Educacion Matemática*, 23(2), 5-32.
- Chapman, L. J. (1967). Illusory correlation in observational report. *Journal of Verbal Learning and Verbal Behavior*, 6(1), 151-155.
- Chapman, L. J. & Chapman, J. P. (1969). Illusory correlation as an obstacle to the use of valid psychodiagnostic signs, *Journal of Abnormal Psychology*, 74, 271-280.
- Contreras, J. M. (2011). *Evaluación de conocimientos y recursos didácticos en la formación de profesores sobre probabilidad condicional*. Tesis doctoral. Universidad de Granada,
- Fint, V. y Godino, J. D. (2006). La noción de configuración epistémica como herramienta de análisis de textos matemáticos: su uso en la formación de profesores. *Educação Matemática Pesquisa* 8 (1), 67-98.

- Godino, J. D. Batanero, C. y Font, V. (2007) The onto-semiotic approach to research in mathematics education. *ZDM. The International Journal on Mathematics Education*, 39 (1-2), 127-135.
- Inhelder, B. y Piaget, J. (1955). *De la logique de l'enfant à la logique de l'adolescent*. (From the child's logic to the adolescent's logic). Paris: Presses Universitaires de France.
- Jenkins, H. M. y Ward, W. C. (1965). Judgment of the contingency between responses and outcomes, *Psychological Monographs*, 79, 1-17.
- Meiser, T. y Hewstone, M. (2006). Illusory and spurious correlations: Distinct phenomena or joint outcomes of exemplar-based category learning? *European Journal of Social Psychology*, 36(3), 315-336.
- Pérez Echeverría, M. P. (1990). *Psicología del razonamiento probabilístico*. Madrid: Ediciones de la Universidad Autónoma de Madrid.
- Shaklee, H. (1983). Human covariation judgment: accuracy and strategy. *Learning and Motivation*, 14, 433-448.
- Shaklee, H. y Mins, M. (1982). Sources of error in judging event covariations: Effects of memory demands, *Journal of Experimental Psychology Learning, Memory and Cognition*, 8(3), 208-224.
- Smedlund, J. (1963). The concept of correlation in adults. *Scandinavian Journal of Psychology*, 4, 165-174.
- Zieffler, A. (2006). *A longitudinal investigation of the development of college students' reasoning about bivariate data during an introductory statistics course*. Unpublished PhD. University of Minnesota.