

# Estrategias en el estudio de la asociación en tablas de contingencia por estudiantes de psicología

Gustavo Cañadas, Carmen Batanero, José Miguel Contreras y Pedro Arteaga

**Resumen:** Las tablas de contingencia son un instrumento necesario en el trabajo de los profesionales de psicología y otras muchas profesiones, aunque se les presta poca atención en la enseñanza universitaria. En este trabajo describimos un estudio exploratorio de las distintas estrategias empleadas para establecer la posible asociación de dos variables presentadas en forma de tabla de contingencia en una muestra de estudiantes de psicología. Se comparan los resultados con un estudio previo de Estepa (1993) con estudiantes de bachillerato y se identifican los conflictos semióticos implícitos en las estrategias parcialmente correctas e incorrectas.

*Palabras clave:* tablas de contingencia, asociación, estrategias, enseñanza universitaria.

## Strategies in the study of the association in contingency tables by psychology students

**Abstract:** Contingency tables are essential tools for the professional work of psychologists and other health-care professionals; however, little attention is paid to its teaching. In this work we describe an exploratory study of the different strategies that a sample of psychology students used to establish the possible association of two variables, presented in contingency tables. Results are compared with a previous study by Estepa (1993) with high school students and implicit semiotic conflicts are identified in the partially correct and incorrect strategies.

*Keywords:* contingency tables, association, strategies, university teaching.

---

Fecha de recepción: 18 de diciembre de 2010.

## INTRODUCCIÓN

Las tablas de doble entrada o tablas de contingencia constituyen un método usual de presentar la información estadística en la prensa o Internet, por lo que algunos autores (por ejemplo, Schield, 2006) incluyen la interpretación correcta de éstas como un componente de la cultura estadística. Estas tablas tienen también interés como instrumento en el ámbito profesional de un psicólogo, especialmente en las actividades de diagnóstico y evaluación psicológica, donde el especialista se enfrenta a diferentes síntomas que pueden estar asociados o no con una patología. Díaz y Gallego (2006) sugieren que, para facilitar el diagnóstico en estas situaciones, se deben organizar los datos en una tabla de contingencia similar a la tabla 1, donde  $A$  representaría una posible patología y  $B$  un síntoma, y la emisión de un diagnóstico sería equivalente a evaluar si existe o no asociación entre las correspondientes variables. A partir de esta tabla, también es posible definir algunos indicadores clínicos habituales que permiten mejorar el diagnóstico, entre otros, la proporción de verdaderos positivos (número de individuos con la enfermedad y el síntoma presente), la proporción de verdaderos negativos (número de individuos sin la enfermedad y el síntoma ausente), sensibilidad y especificidad de la prueba.

**Tabla 1.** Tabla de contingencia  $2 \times 2$

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

Sería importante conseguir que los estudiantes que salen de las facultades de psicología interpreten correctamente las tablas de contingencia, tanto las similares a la tabla 1 como otras de mayor número de filas y columnas, para poder calcular e interpretar los índices mencionados, así como otros que los ayuden a tomar decisiones en su vida profesional. A pesar de esta necesidad educativa, dichas tablas reciben poca atención en la enseñanza universitaria, pues se supone que su lectura e interpretación son habilidades adquiridas por estos estudiantes.

El objetivo de este trabajo es realizar un estudio exploratorio de evaluación de la comprensión intuitiva de la asociación entre las dos variables presentes en una tabla de contingencia por parte de los estudiantes que ingresan en Psicología. Más concretamente, pretendemos analizar las estrategias que estos alumnos utilizan en el análisis de la asociación en dichas tablas y establecer una comparación con los resultados de estudios previos con otros tipos de estudiantes. Los resultados se utilizarían para organizar acciones formativas dirigidas a mejorar esta comprensión.

## MARCO TEÓRICO

Siguiendo a Font, Godino y D'Amore (2007), en nuestro trabajo asumimos que, en las prácticas matemáticas que realizan los sujetos al resolver problemas, se presentan múltiples funciones semióticas (o bien de lectura, o bien de representación), debido a la necesidad de usar y operar con objetos matemáticos que son inmateriales. Los autores utilizan la idea de función semiótica en el sentido de Eco (1977), quien las define como correspondencia entre un antecedente (expresión) y un consecuente (contenido) establecida por un sujeto. Más concretamente, el trabajo se centra en las prácticas matemáticas involucradas en la realización de un juicio de asociación sobre variables presentadas en una tabla de contingencia. Observamos que, en esta tarea, el estudiante debe realizar varias actividades de traducción entre la tabla en su conjunto o una parte de ella y lo representado.

Nuestro supuesto es que la tabla de contingencia, a pesar de su simplicidad aparente, puede considerarse como un objeto semiótico complejo. Al igual que lo indicado por Bertin (1967) para la lectura de gráficos, la tabla es un texto multimodal; tanto en su conjunto como por los elementos que lo componen, que están constituidos por conjuntos de signos que requieren una actividad semiótica de quienes los interpretan. La lectura de la tabla comienza con una identificación externa del tema al que se refiere a través de la lectura del enunciado del problema. A continuación se requiere una identificación interna de las dimensiones relevantes de variación en la tabla, es decir, la interpretación de las variables representadas y sus valores. Finalmente se produce una percepción de la correspondencia entre las frecuencias de los diferentes valores de cada variable para obtener conclusiones sobre su posible asociación y sus relaciones con la realidad representada.

Por otro lado, Font, Godino y D'Amore (2007) consideran una tipología de objetos matemáticos (expresiones verbales o simbólicas, propiedades, procedimientos, problemas, argumentos, conceptos), que intervienen en las prácticas matemáticas y donde cada una de los cuales puede desempeñar el papel de antecedente o consecuente de una función semiótica. Las prácticas realizadas por los participantes en el estudio los llevan a operar con una diversidad de objetos, por ejemplo, frecuencias absolutas y relativas dobles, condicionales o marginales, independencia y asociación. Cada uno de estos objetos, tanto en el enunciado de los problemas como en las soluciones de los estudiantes, se expresa por escrito, lo que lleva a establecer una función semiótica, cuya correspondencia entre antecedente y consecuente suele estar implícita.

Si las interpretaciones realizadas por los alumnos no son las esperadas por el profesor, se produce un *conflicto semiótico*. Los autores denominan conflicto semiótico a cualquier disparidad o discordancia entre los significados atribuidos a una expresión por dos sujetos, en este caso, profesor y alumno.

Nuestro trabajo se orienta al estudio de las prácticas matemáticas realizadas por los estudiantes para establecer un juicio de asociación en una tabla de contingencia. Dichas prácticas serán clasificadas en estrategias diferenciadas, cuyo análisis permitirá la identificación de posibles conflictos semióticos de los estudiantes. En lo que sigue se describen las investigaciones previas, el método utilizado y se discuten los resultados obtenidos.

## INVESTIGACIONES PREVIAS

La investigación sobre el estudio de la asociación en tablas de contingencia es escasa en educación matemática, aunque tiene una mayor tradición en el campo de la psicología. En lo que sigue, analizamos estos dos tipos de investigaciones.

### INVESTIGACIONES EN PSICOLOGÍA

En el campo de la psicología, la investigación sobre comprensión de las tablas de contingencia la inician Inhelder y Piaget (1955) como parte de su estudio del desarrollo evolutivo del razonamiento sobre probabilidad. Como los autores pensaban que la comprensión de la asociación se adquiere sólo cuando se ha conseguido la comprensión del concepto de probabilidad, realizaron su estudio

con estudiantes a partir de los 12 años. Se interesaron por las estrategias en los juicios de asociación, usando como material experimental tarjetas con dibujos de caras en las que se presentaban todas las combinaciones posibles de color de pelo (rubio o moreno) y color de ojos (azules o negros). Para explicar sus resultados, los autores utilizaron una tabla similar a la tabla 1, donde  $A$  es ahora el color de pelo rubio y  $B$  el color de ojos azules.

Los resultados de estas investigaciones muestran que la mayoría de los chicos de 12-13 años sólo utilizan la celda  $a$  (donde se dan simultáneamente la presencia de los caracteres  $A$  y  $B$ ; en el ejemplo, individuos con ojos azules y pelo rubio) para juzgar la asociación, sin entender que la celda  $d$  (ausencia simultánea de los dos caracteres, en el ejemplo, sujetos con pelo oscuro y ojos oscuros) tiene el mismo peso en relación con la asociación.

En una segunda etapa, que aparece entre los 13-15 años, los chicos solamente comparan celdas dos a dos (por ejemplo comparan  $a$  con  $b$ , es decir, el número de sujetos con ojos azules entre los que tienen el pelo rubio o moreno). Otro nivel posterior sería comprender cuales son los casos favorables ( $a$  y  $d$ ) y desfavorables ( $b$  y  $c$ ) de la asociación sin compararlos. Finalmente, se establecen las relaciones diagonales, comparándolas entre sí o con el total ( $a + b + c + d$ ).

Como consecuencia del trabajo de Inhelder y Piaget, diversos autores, como Smedlund (1963) y Shaklee y Mins (1982), han tratado de comprobar si los sujetos adultos llegan a la estrategia considerada correcta por Piaget de una manera natural en su desarrollo. En contraposición con lo supuesto por Piaget e Inhelder, se observa que hay adultos que continúan con estrategias propias de niños en el periodo preoperatorio, es decir, que al realizar un juicio de asociación, sólo se fijan en la celda  $a$  o bien usan estrategias de la segunda etapa (comparando la celda  $a$  con la  $b$  o la  $d$ ). En consecuencia, los sujetos adultos no siempre utilizan todos los datos relevantes en el problema y el desarrollo del concepto de asociación no sería espontáneo en los sujetos, a menos que se dé una instrucción formal sobre él.

Jenkins y Ward (1965) muestran la dificultad de los juicios de asociación, señalando que incluso la estrategia de comparar las diagonales, considerada como correcta por Inhelder y Piaget, sólo llevaría a un juicio de asociación correcto si las frecuencias marginales por filas o por columnas en la tabla 1 fuesen iguales (en el ejemplo, si hubiese igual número de sujetos rubios y morenos o de sujetos con ojos azules y oscuros). En otros casos, puede dar lugar a un juicio de asociación incorrecto, aunque a pesar de ello, esta estrategia es utilizada por adultos en las investigaciones de Shaklee y Tucker (1980) y Allan y Jenkins (1983).

Jenkins y Ward indican que la estrategia correcta para el caso general consiste en comparar la diferencia entre las probabilidades  $P(B|A)$  y  $P(B|\text{no } A)$  de la manera siguiente:

$$d = \left( \frac{a}{a+c} \right) - \left( \frac{b}{b+d} \right)$$

Pérez Echeverría (1990) realiza una compilación de las investigaciones anteriores y otras investigaciones en psicología sobre estrategias en los juicios de asociación y las clasifica en niveles de dificultad de la manera siguiente, señalando que lo más frecuente es el nivel 3 o 4 y rara vez se llega al último nivel.

- Nivel 1. Utilizar sólo una celda, generalmente la celda a.
- Nivel 2. Comparar la celda a con b o c, es decir usar dos celdas.
- Nivel 3. Comparar la celda a con b y c, es decir usar tres celdas.
- Nivel 4. Usar todos los datos pero sólo mediante comparaciones aditivas.
- Nivel 5. Concluir la asociación por el uso de una relación multiplicativa entre las cuatro celdas.

### EFECTO DE LAS TEORÍAS PREVIAS

Como hemos visto, la realización de juicios de asociación no es una tarea sencilla, pues hay estrategias incorrectas que llevan a un error en el juicio. Para agravar el problema, Chapman y Chapman (1969) mostraron que muchas personas forman sus propias teorías sobre la relación entre variables en la tabla de contingencia, lo que les impide evaluar correctamente la asociación. La experiencia vital y el entorno cultural son los responsables de estas teorías, sobre todo en relación con variables que llaman la atención de las personas. Un ejemplo sería el encontrar una falsa asociación entre la pertenencia a un grupo minoritario y los comportamientos del grupo.

Este fenómeno se conoce como “correlación ilusoria” y lleva a la percepción de una relación donde no existe ninguna, o bien a la percepción de una relación más fuerte de la que en realidad existe (Jennings, Amabile y Ross, 1982; Alloy y Tabacnick, 1984; Wrigth y Murphy, 1984).

## INVESTIGACIÓN EN DIDÁCTICA DE LA MATEMÁTICA

En Didáctica de la Matemática este tema no ha recibido demasiada atención. Una excepción son los trabajos de Estepa (1993) y Estepa y Batanero (1995), quienes realizan un estudio cualitativo de estrategias de los 213 estudiantes del curso preuniversitario y describen las siguientes categorías:

- Estrategias correctas. (E1) comparar todas las distribuciones de frecuencias relativas condicionales de una variable para los distintos valores de la otra variable; (E2) comparar todas las frecuencias relativas condicionadas de una variable para un único valor de la otra variable con la frecuencia marginal de la primera variable y (E3) comparación de posibilidades a favor y en contra de B en cada valor de A.
- Estrategias parcialmente correctas. (E4) comparar la distribución de frecuencias absolutas dobles de una variable para un único valor de la otra variable con la frecuencia marginal de la primera variable; (E5) comparar solamente una de las frecuencias absolutas dobles de una variable para un valor de la segunda con su frecuencia marginal de la primera variable; (E6) comparar las frecuencias absolutas dobles entre sí; (E7) comparar la suma de frecuencias en las diagonales.
- Estrategias incorrectas. (E8) usar sólo la celda de mayor frecuencia; (E9) usar sólo una distribución condicional; (E10) comparar frecuencias conjuntas con el número total de observaciones y (E11) comparar frecuencias marginales entre sí.

Estepa (1993) sugiere que las estrategias incorrectas usadas por los estudiantes en los juicios de asociación dependen de algunas concepciones que poseen los alumnos sobre ésta y describe las siguientes:

- Concepción causal. El sujeto sólo considera la dependencia entre variables si puede explicarse por la presencia de una relación causal entre ellas.
- Concepción determinista. Son los sujetos que sólo perciben la dependencia de tipo funcional y no son capaces de generalizar al caso de dependencia aleatoria. Un ejemplo es cuando se exige la existencia de una expresión algebraica que relacione las variables.
- Concepción unidireccional. En este caso el estudiante no admite la asociación inversa, considera la intensidad de la asociación pero no su signo.

- Concepción local. Esta concepción se presenta cuando los sujetos dan su solución basándose únicamente en algunos casos aislados, por ejemplo, si sólo se considera una distribución condicional o la celda de máxima frecuencia.

El objetivo de este trabajo es analizar hasta qué punto las estrategias y concepciones descritas por Estepa en estudiantes de bachillerato se reproducen en los alumnos que ingresan en psicología, pues estas estrategias y concepciones incorrectas pueden incidir en el futuro en errores en la práctica del diagnóstico psicológico. A continuación se describen el método y los resultados.

## MÉTODO

La muestra estuvo formada por 62 alumnos de primer año de la Licenciatura en Psicología de la Universidad de Huelva, que cursaban una asignatura de Análisis de Datos I. Los datos se tomaron antes del estudio específico de las tablas de contingencia, a fin de detectar posibles estrategias y concepciones intuitivas incorrectas y tenerlas en cuenta en el diseño de la enseñanza.

El cuestionario que se usó está adaptado del de Estepa y Batanero (1995) y se compone de cuatro ítems de respuesta abierta con la misma presentación de los datos y preguntas realizadas, aunque variando el contexto: padecer estrés o no y tener insomnio o no (figura 1); ser o no hijo único y ser o no un niño problemático (figura 2); llevar o no una vida sedentaria y tener o no alergia (figura 3) y número de horas dedicadas a estudiar un examen y aprobar o suspender (figura 4, donde el número de horas tiene 3 categorías). Las frecuencias de todas las celdas fueron números menores que 100 y la frecuencia total estuvo comprendida entre 100 y 250 casos. Comparado con el estudio de Estepa, se aumentó la frecuencia en los datos y se cambió el contexto a otros relacionados con el campo de la psicología en los ítems 1 y 2.

Una solución correcta a estos ítems consistiría en comparar las distribuciones de frecuencias relativas condicionales. Por ejemplo, en el ítem 1, la frecuencia relativa padecer estrés cuando se tiene insomnio es  $\frac{90}{150} = 0.6$  y la frecuencia relativa de padecerlo si no se tiene  $\frac{60}{100} = 0.6$ , que son iguales. Asimismo, son iguales las frecuencias relativas de no padecer estrés si se tiene insomnio (0.4) y



**Figura 1. Ítem 1**

Se pretende estudiar si el sufrir insomnio tiene relación con los trastornos de estrés. En una muestra de 250 personas observadas, se obtuvieron los siguientes resultados:

	Padece estrés	No padece estrés
Tener insomnio	90	60
No tener insomnio	60	40

- Utilizando estos datos, razona si tener o no insomnio tiene relación con padecer estrés.
- Indica cómo has usado los datos de la tabla para llegar a tu conclusión

**Figura 2. Ítem 2**

Un psicólogo infantil estudia si existe relación entre ser hijo único (sin hermanos) y ser un niño problemático. Para ello se han observado 250 niños con los resultados siguientes:

	Niño problemático	Niño no problemático
Tiene hermanos	40	100
Hijo único	100	10

- Utilizando estos datos, razona si hay alguna relación entre ser un niño problemático y tener o no hermanos.
- Indica cómo has utilizado los números de la tabla para llegar a tu conclusión.

Figura 3. Ítem 3

Se quiere saber si sufrir o no de alergia tiene relación con llevar una vida sedentaria (llevar una vida sin realizar ningún tipo de ejercicio físico). Para ello, obtuvimos los siguientes resultados de 300 sujetos:

	Sufre alergia	No sufre alergia
Forma de vida sedentaria	130	30
Forma de vida no sedentaria	20	120

- Utilizando estos datos, razona si padecer alergia tiene o no relación con llevar una vida sedentaria en estos sujetos.
- Indica cómo has usado los números de la tabla para obtener tu conclusión.

Figura 4. Ítem 4

La siguiente tabla nos muestra el número de estudiantes que aprobaron o suspendieron un examen teniendo en cuenta el tiempo que cada estudiante dedicó a prepararlo.

	Menos de 5 h	Entre 5 y 10 h	Más de 10 h	Total
Aprobados	5	15	51	71
Suspensos	20	7	2	29
Total	25	22	53	100

- Utilizando estos datos, razona si aprobar o suspender el examen tiene relación con el tiempo dedicado a prepararlo.
- Indica cómo has usado los números de la tabla para llegar a tu conclusión.

si no se tiene (0.4). Por tanto, las variables en este ejemplo son independientes, como se indica en la tabla 2.

Las variables de tarea que se tuvieron en cuenta en el momento de la elaboración del cuestionario (véase tabla 2) fueron las siguientes:

1. *Signo de la asociación entre las variables*, considerándose los tres casos posibles: dependencia directa, dependencia inversa e independencia. En el caso del ítem 1, al ser independientes las variables, el valor del coeficiente es igual a cero.
2. *Intensidad de la dependencia*, medida mediante el coeficiente Phi de Pearson en tablas  $2 \times 2$  y con el coeficiente V de Cramér en tablas  $2 \times 3$ . Se eligieron un ítem de intensidad cero (ítem 1), otro de intensidad moderada-baja (ítem 4) y dos de intensidad moderada-alta (ítems 2 y 3).
3. *Concordancia entre los datos y las teorías previas sugeridas por el contexto*. Se usaron ítems en los que coincide la asociación empírica en los datos con las teorías previas (ítems 2 y 4), otros donde no coinciden (ítem 1) y un ejemplo donde no hay teorías previas (ítem 3).
4. *Tipo de covariación*. Utilizamos tres categorías de Barbancho (1973): dependencia causal unilateral (ítems 2 y 4), interdependencia (ítem 1), dependencia indirecta (ítem 3).

**Tabla 2.** Variables de tarea en los ítems

	Tabla $2 \times 2$			Tabla $2 \times 3$
	ítem 1	ítem 2	ítem 3	ítem 4
Dependencia	Independencia	Inversa	Directa	Directa
Valor del coeficiente de asociación	0	-0.62	0.67	0.37
Concuenda con teoría previa	No	Sí	No hay teoría	Sí
Tipo covariación	Interdependencia	Causal unilateral	Dependencia indirecta	Causal unilateral

## RESULTADOS Y DISCUSIÓN

Una vez recogidos los cuestionarios, se realizó un análisis cualitativo de su contenido, teniendo en cuenta el juicio de asociación dado en la pregunta *a* del ítem, donde el alumno indica si piensa que las variables están o no relacionadas y la estrategia utilizada (que el alumno describe en el apartado *b* del ítem). Siguiendo la investigación de Estepa (1993), se clasificaron las estrategias en primer lugar en correctas, parcialmente correctas e incorrectas y, dentro de ellas, se diferenciaron en función de los objetos matemáticos usados y los conflictos semióticos presentados. A continuación se describe la clasificación de estrategias mostrando un ejemplo de cada una de ellas.

*Estrategias correctas.* Se basan en el uso de todos los datos de la tabla mediante la comparación de probabilidades, es decir, estarían en el nivel 5 de Pérez Echevarría (1990) y se han encontrado las siguientes:

SC.1. *Comparar todas las distribuciones de frecuencias relativas condicionales de una variable para los distintos valores de la otra variable*, estrategia descrita por Estepa (1993). En el ejemplo que se presenta a continuación, el alumno compara en el ítem 1 el porcentaje de personas con estrés en el grupo de personas con y sin insomnio. Como el porcentaje es el mismo en ambos casos (60%), deduce la independencia de las variables, utilizando implícitamente la siguiente propiedad: la independencia de una variable *B* respecto a otra variable *A* implicaría que las frecuencias relativas condicionales  $h(B|A)$  cuando *A* cambia no varían.

“Si observamos esta tabla se ve claramente que 60% de las personas sin insomnio padece estrés, por lo que resulta una regla de tres:

$$\left( \begin{array}{l} 90 \rightarrow 150 \\ ? \rightarrow 100 \end{array} \right) = \frac{90 \times 100}{150} = 60$$

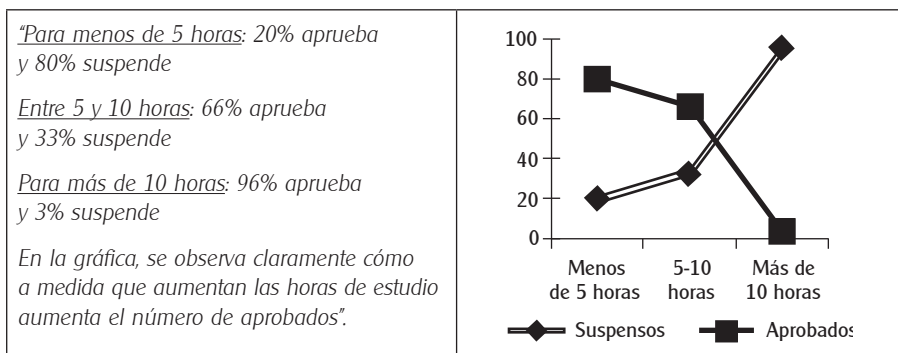
*Se observa que 90 es 60% de 150, es decir, hay el mismo porcentaje de personas con estrés, con o sin insomnio”.*

SC.2. *Comparación de posibilidades a favor y en contra de B en cada valor de A*, también descrita por Estepa (1993). Un ejemplo en el ítem 2 se presenta a continuación: “Operamos sobre 140 niños con hermanos, 40 con problemas y 100 sin ellos. Simplificando, podemos decir que 4 de cada 14 niños con

hermanos presentarán problemas. También podemos afirmar que 10 de cada 11 hijos únicos presentan problemas". El alumno emplearía implícitamente la propiedad de que existe una correspondencia unívoca entre la probabilidad de un suceso y la razón de sus probabilidades a favor y en contra.

SC3. El alumno calcula las frecuencias relativas condicionales por filas y las representa gráficamente. Utiliza el crecimiento o decrecimiento para deducir la asociación. Un ejemplo en el ítem 4, donde el alumno considera que si hay relación se transcribe en la figura 5. Observamos que, como en la estrategia SC1 el alumno compara las diferentes distribuciones condicionales. La diferencia es que lo hace gráficamente y usa las propiedades gráficas para deducir la asociación. Esta estrategia no fue descrita por Estepa (1993).

Figura 5. Ejemplo de la estrategia correcta SC3



*Estrategias parcialmente correctas.* Son estrategias que usan todos los datos de la tabla, pero se basan en la comparación de frecuencias absolutas en vez de comparar probabilidades. Por lo general se basan en comparaciones aditivas y estarían en el nivel 4 de Pérez Echevarría (1990). Se han encontrado las siguientes:

SP.1. Comparar la distribución de frecuencias absolutas condicionales con la frecuencia absoluta marginal correspondiente, utilizando comparaciones aditivas o cualitativas de tal modo que no se cuantifican bien las probabilidades. Un ejemplo en el ítem 1 se incluye a continuación: "Observando la tabla, se deduce que 150 personas tienen insomnio, y de ellos 90 padecen estrés; además 100 personas no lo tienen y sólo 60 de ellos padecen estrés. Por lo que se puede deducir, según las probabilidades, que sí puede tener relación". Aunque

en caso de dependencia las distribuciones condicionales serán diferentes, esta respuesta (al contrario de la estrategia ST.1) se basa en frecuencias absolutas y no relativas. La respuesta muestra diferentes conflictos. Por un lado, en vez de comparar probabilidades, compara sólo los casos favorables, lo que implica una posible confusión entre probabilidades y frecuencias. Por otro lado, el estudiante no es consciente de que muestras de diferentes tamaños no pueden compararse mediante frecuencias absolutas.

SP.2. *Comparar las frecuencias absolutas condicionadas entre sí.* Esta estrategia es parecida a la SC.1, aunque comparando las frecuencias absolutas entre sí en lugar de frecuencias relativas. Un ejemplo en el ítem 4 es el siguiente: “*A medida que va aumentando el tiempo, los aprobados van creciendo y los suspensos van disminuyendo. Hay una relación causal directa*”. El alumno hace una comparación de las frecuencias absolutas en la fila de número de aprobados y, aunque detecta el crecimiento de las frecuencias, ello solo no basta para deducir la dependencia, pues sería posible que también creciese el número de suspensos en función de las horas de estudio. El conflicto en este estudiante se produce porque usa frecuencias absolutas en lugar de relativas. Además, muestra una confusión entre causalidad y asociación.

SP.3. *Comparar la suma de frecuencias en las diagonales de la tabla,* estrategia descrita por Inhelder y Piaget (1955) que es correcta sólo en el caso de que la tabla tenga iguales frecuencias marginales para la variable independiente, pero no en nuestro estudio. Un ejemplo en que el alumno considera que sí hay relación en el ítem 2 es el siguiente: “*Simplemente he mirado los resultados de la investigación, viendo que entre los hijos únicos y problemáticos y los niños que tienen hermanos y no son problemáticos se suman prácticamente todos los sujetos estudiados, quedando sólo 50 de ellos con las características opuestas a las antes descritas*”. El conflicto aparece al ser diferentes las frecuencias marginales, lo que no es percibido por el alumno.

SP.4. *Equiparar la independencia con equiprobabilidad en las celdas de la tabla,* estrategia que no hemos encontrado descrita en las investigaciones previas. El alumno supone que, para que se dé la independencia, todas las frecuencias relativas dobles de las celdas de la tabla deben ser iguales (es decir 25% de casos en cada celda). Esta afirmación sólo sería cierta en el caso de que las frecuencias marginales por filas y columnas fuesen iguales en la tabla, pero no en el caso general. Un ejemplo en el ítem 1, donde el alumno argumenta que existe asociación, se presenta a continuación: “*He dividido 250 entre 4 (25%) para ver el número de personas que tendría que haber en cada cuadrado*”

(25%) si no afectara tener insomnio. Tras esto, he visto que las personas que no tienen insomnio y padecen estrés están cerca de 25%, igual que las personas que tienen insomnio y no padecen estrés. Las personas que no tienen insomnio ni estrés son menos de 25% y los que tienen insomnio y estrés sobrepasan mucho 25%".

SP.5. Comparación de frecuencias de casos a favor y en contra de  $B$  en un único valor de  $A$  que fue descrita por Estepa (1993), aunque nosotros la extendemos también al caso de *error al compararlas en todos los valores*. El conflicto se produce porque el cociente de frecuencias de casos a favor y en contra (riesgo relativo) se calcula incorrectamente, pues debiera haberse calculado el riesgo de  $B$  en las dos distribuciones condicionales,  $A$  y no  $A$ . En lugar de ello, se compara el riesgo de  $B$  si se da no  $A$  con el riesgo de  $A$  si se da  $B$ . Un ejemplo en el ítem 1, donde el alumno argumenta que existe asociación, se presenta a continuación: "Se podría calcular el estrés como factor de riesgo a través de la fórmula del riesgo relativo o OR. Si es  $>1$  el estrés es un factor de riesgo. Al ser  $>1$ , el estrés es un factor de riesgo para padecer insomnio".

$$OR = \frac{\frac{a}{a+b}}{\frac{b}{b+c}} \Rightarrow \frac{\frac{90}{90+60}}{\frac{60}{60+60}} = \frac{0.6}{0.5} = 1.7$$

*Estrategias incorrectas.* Cuando el estudiante usa sólo una parte de los datos necesarios, por tanto estaría en uno de los niveles 1 a 3 de Pérez Echevarría (1990). Se han encontrado las siguientes.

SI.1. *El uso único de la celda de mayor frecuencia.* En este caso el alumno no utiliza toda la información presente en el problema, sino que se conforma con juzgar la asociación a partir de la celda más sobresaliente. Estos alumnos mostrarían la concepción local de la asociación descrita por Estepa (1993), pues tratan de decidir sobre la asociación sin tener en cuenta todos los datos del problema. Un ejemplo en que el alumno indicó que existía relación en el ítem 4 es: "100 personas: 51 aprobaron con más de 10 horas de estudio". En esta respuesta, al conflicto de tener en cuenta sólo una parte de los datos se añade el utilizar frecuencias absolutas en lugar de relativas.

SI.2. *Uso de sólo una distribución condicional para deducir la asociación,* descrita por Estepa (1993). En este caso el alumno no ve el problema como un problema de comparación de probabilidades y muestra una concepción local de la asociación. Un ejemplo donde el alumno justifica la existencia de asociación

en el ítem 4 se presenta a continuación: *“Simplemente he observado los datos de la tabla y, según más horas se les dedique de estudio, existe un mayor número de aprobados”*. Este alumno sólo usa los datos de los alumnos aprobados, estudiando el crecimiento o decrecimiento. Además utiliza frecuencias absolutas en vez de relativas.

SI.3. *Comparar frecuencias conjuntas absolutas o relativas entre sí o con el número total de observaciones*. Un ejemplo en el ítem 1 para justificar la asociación es el siguiente:

*“He usado los datos de cada variable y he hecho el % de cada uno de ellos respecto al total; y he comparado los resultados; como son diferentes, están relacionadas las variables:*

$$\frac{90}{250} \times 100 = 36\%$$

$$\frac{60}{250} \times 100 = 24\%$$

$$\frac{60}{250} \times 100 = 24\%$$

$$\frac{40}{250} \times 100 = 16\%”$$

Por un lado, el alumno asume incorrectamente que, para que se dé la independencia, los porcentajes en todas las frecuencias conjuntas deberían ser iguales, por tanto, aplica incorrectamente una propiedad que es cierta para las frecuencias condicionales al caso de las frecuencias conjuntas mediante una generalización abusiva. El conflicto se produce al no diferenciar entre frecuencias conjuntas y condicionales. El error es similar al descrito por Falk (1986) en relación con la probabilidad condicional, donde los alumnos confunden probabilidad condicional y conjunta. Por otro lado, aunque en este caso el alumno calcula en primer lugar las frecuencias relativas dobles de todas las celdas, para realizar el juicio de asociación, elige entre ellas únicamente la celda de mayor frecuencia, por lo que manifiesta la concepción local de la asociación (Estepa, 1993).

SI.4. *Comparar frecuencias marginales entre sí*, asumiendo incorrectamente que, al obtener distintas frecuencias marginales, los datos no se pueden comparar, error advertido por Watson (2006) en relación con la comparación de distribuciones. Un ejemplo en el ítem 2 sería el siguiente: *“Al calcular la media como hice en el ejercicio anterior, he comprobado que hay más niños con problemas (70) que niños sin problemas (55), no se puede comparar”*. Por un lado, aparece una confusión entre frecuencias y valor de la variable, ya que la media se calcula



sobre los valores y no sobre las frecuencias. Por otro lado, trataría de comparar entre sí las dos medias de una distribución marginal.

Las estrategias anteriores habían sido descritas por Estepa o en algún caso por otros autores. También hemos identificado algunas estrategias no descritas en las investigaciones previas, son las siguientes.

SI.5. *Usa sólo la celda en que los dos caracteres están ausentes y, si la frecuencia es alta, supone que esto implica la no existencia de asociación.* Aunque en el apéndice de su tesis Estepa (1993) diferencia el uso de las celdas *b*, *c* y *d* como estrategias diferenciadas, no hace mención explícita a la búsqueda de la celda *d* (ausencia-ausencia) que en realidad es favorable a la asociación. Es una estrategia incorrecta, pues se basa en frecuencias absolutas. Un ejemplo en el ítem 3 para justificar que no hay asociación es el siguiente: “*No hay relación, porque hay bastantes sujetos que no llevan una vida sedentaria y no sufren alergia*”. El conflicto aparece al no considerar la celda *d* como favorable a la asociación, además de usar frecuencias absolutas.

SI.6. *Compara la casilla a con el resto de las celdas, o bien con las celdas b y c, esperando que estas celdas estén en contra de la asociación.* Es decir, comparan los casos en que se presentan los dos caracteres conjuntamente con aquellos en que se presenta sólo uno de los dos o ninguno de los dos. Es una variante de la estrategia “usar la casilla *a*” descrita por Estepa (1993). En el siguiente ejemplo (ítem 3), el alumno compara la celda vida sedentaria-sufrir alergia, con la no sedentaria-no sufrir alergia, observando que los valores son casi iguales. El conflicto aparece al asumir incorrectamente que esta igualdad implica independencia. Un ejemplo aparece a continuación: “*Fácilmente se observa que el unir las dos variables “forma de vida sedentaria-sufrir alergia” tiene una cifra elevadísima comparada con la unión de variables forma de vida sedentaria-no sufre alergia. También se puede apreciar que llega casi a igualarse a la cifra perteneciente a la unión de variables forma de vida no sedentaria-no sufre alergia*”.

SI.7. *El uso de las celdas de mayor y menor frecuencia.* En este caso, el alumno no utiliza toda la información presente en el problema, tan sólo las celdas más sobresalientes. Es una variante de la descrita por Estepa (1993) que consiste en fijarse únicamente en la celda de mayor frecuencia. Un ejemplo en el ítem 4 sería el siguiente: “*Sólo 5 personas de 100 aprobaron un examen estudiando menos de 5 horas frente a las 51 personas de 100 que aprobaron el examen estudiando más de 10 horas*”. El alumno asume que el crecimiento observado implica diferencia. Aunque el alumno llega a detectar la asociación en los datos, esta estrategia para un caso general sería incorrecta pues, además de usar parte de los datos, se

basa en frecuencias absolutas.

SI.8. *El uso de las celdas de menor frecuencia.* Como en el caso anterior, el alumno no utiliza toda la información presente en el problema, tan sólo la celda más sobresaliente. Por tanto, muestra la concepción local de la asociación. Un ejemplo en el ítem 4 sería el siguiente: “*Si, ya que, según la tabla, a más de 10 h de estudio menos probabilidad de suspender sólo 2 de 53*”. El conflicto aparece al no usar todos los datos del enunciado.

SI.9. *El alumno forma las tablas de frecuencias marginales,* calculando frecuencias, probabilidades, porcentajes. Es una estrategia incorrecta y no lleva a finalizar el juicio de asociación. Un ejemplo en el ítem 4 para justificar la asociación se presenta en la figura 6.

**Figura 6.** Ejemplo de la estrategia incorrecta SI.10

Tabla de frecuencia (aprobados)					Tabla de frecuencia (suspensos)					
x	$f_i$	$p_i$	$fa$	$P_1$		x	$f_i$	$p_i$	$fa$	$P_1$
5H	5	0.07	5	7		5H	20	0.68	20	68
5-10 H	15	0.23	20	21		5-10 H	7	0.24	27	24
+10 H	51	0.71	71	71		+10 H	2	0.07	29	7
	N=71	1		00%			N=25	1		99%

SI.10. *Calcula la diferencia de las diagonales.* Estepa (1993) describe la comparación de la suma de diagonales, que sería una estrategia parcialmente correcta. Un ejemplo en el ítem 3 para justificar la no asociación es: “*Observando la poca diferencia entre 130-120 y 20-30 (sólo 10)*”. El estudiante calcula la diferencia de frecuencias en la diagonal que apoya la relación y la que la contradice. El conflicto aparece al asumir incorrectamente que esta pequeña diferencia es indicativa de no asociación, en lugar de comparar la suma de las diagonales.

SI.11. El alumno sólo considera sus teorías previas y no tiene en cuenta los datos. Muestra el sesgo denominado correlación ilusoria que lleva a la percepción de una relación donde no existe ninguna, o bien a la percepción de una relación más fuerte de la que existe en realidad (Jennings, Amabile y Ross, 1982; Wrigth y Murphy, 1984; Alloy y Tabacnick, 1984). No utiliza los datos del problema ni ningún otro objeto matemático, por lo que no realizamos el análisis semiótico. Es incorrecta. Un ejemplo en el ítem 1 sería el siguiente: “*porque la*

*variable de tener insomnio conlleva a padecer estrés, y no tener insomnio no conlleva a tener estrés”.*

SI.12. *Otros procedimientos incorrectos.* En casos esporádicos, se usan procedimientos no relacionados con las frecuencias de la tabla, por ejemplo, plantear una ecuación, o bien no se especifica con claridad el procedimiento. Un ejemplo en el ítem 2 sería el siguiente: *“Por la observación de la tabla y utilizando la lógica entre los datos que se estudian”.*

En la tabla 3 se presentan las frecuencias de cada una de estas estrategias. Las estrategias más frecuentes fueron SP2 (comparar frecuencias absolutas condicionales entre sí), SI2 (uso de una sola distribución condicional) y SI3 (Comparar frecuencias conjuntas absolutas o relativas entre sí. Hay coincidencia con la investigación de Estepa (1993) en que la estrategia más frecuente fue utilizar una única distribución condicional. Sin embargo, otras dos estrategias frecuentes en el estudio de Estepa fueron comparar posibilidades o razón (SC2) y comparar frecuencias marginales (SI4) que han tenido menor uso en nuestro estudio.

En la tabla 4 se resumen las estrategias y una comparación con los resultados de Estepa (1993) en ítems similares. Observamos cerca de 60% de estrategias incorrectas, disminuyendo algo el porcentaje en el ítem 4, donde algunos alumnos han usado el crecimiento/decrecimiento de los datos (estrategia SP2, usada por 30%, que es parcialmente correcta). Además el contexto es familiar y la teoría sobre la asociación se confirma en los datos. De todos modos, el porcentaje de estrategias incorrectas es aún muy elevado. Suben en el ítem el porcentaje de estrategias parcialmente correctas respecto al resto, donde se presentan en alrededor de 20% de los casos y las estrategias correctas varían entre 8 y 15% según el ítem, proporción claramente insuficiente tratándose de alumnos universitarios que tendrán que hacer uso de las tablas de contingencia en su profesión.

Al comparar estos resultados con los obtenidos por Estepa (1993), se puede apreciar que el porcentaje de estrategias correctas en todos los casos es muy similar en las dos muestras, a pesar de que han variado algunas de las estrategias respecto a las mencionadas por el autor. Por el contrario, el porcentaje de estrategias parcialmente correctas fue bastante mayor en la investigación de Estepa, excepto en el primer ítem correspondiente al caso de independencia en donde muchos de sus estudiantes dejaron la respuesta en blanco. Nuestros estudiantes han dado mayor número de estrategias incorrectas en todos los ítems, aunque dejan menos respuestas en blanco.

En la tabla 5, se hace un resumen de los conflictos semióticos detectados en las estrategias parcialmente correctas e incorrectas. Los hemos clasificado según

Tabla 3. Frecuencias de estrategias en los diferentes ítems

(y porcentajes)		Ítem 1	Ítem 2	Ítem 3	Ítem 4
Correctas	SC1	6(9.7)	6(9.7)	6(9.7)	5(8.1)
	SC2	3(4.8)	1(1.6)		1(1.6)
	SC3*	3(4.8)	5(8.1)	5(8.1)	5(8.1)
Parcialmente correctas	SP1	4(6.4)		2(3.2)	1(1.6)
	SP2	8(12.9)	7(11.3)	6(9.7)	19(30.6)
	SP3		7(11.3)	5(8.1)	2(3.2)
	SP4*	1(1.6)			
	SP5*	1(1.6)			
Incorrectas	SI1	8(12.9)	4(6.5)	6(9.7)	4(6.5)
	SI2	7(11.3)	11(17.7)	9(14.5)	10(16.1)
	SI3	8(12.9)	12(19.4)	11(17.7)	4(6.5)
	SI4	3(4.8)	2(3.2)	2(3.2)	1(1.6)
	SI5*	1(1.6)		2(3.2)	
	SI6*	3(4.8)		1(1.6)	
	SI7*				1(1.6)
	SI8*		1(1.6)		1(1.6)
	SI9*				1(1.6)
	SI10*			1(1.6)	
	SI11*	2(3.2)	1(1.6)	1(1.6)	
	SI12	2(3.2)	3(4.8)	3(4.8)	2(3.2)
	No responde	2(3.2)	2(3.2)	2(3.2)	5(8.1)
Total	62	62	62	62	

\* Estrategias no descritas por Estepa.

**Tabla 4.** Porcentaje de estrategias en la investigación actual y en Estepa (1993)

	Correctas		Parcialmente correctas		Incorrectas		Blanco	
	Actual	Estepa	Actual	Estepa	Actual	Estepa	Actual	Estepa
Ítem 1	19.4	19.3	22.6	10.8	54.8	50.2	3.2	19.7
Ítem 2	19.4	20.2	22.6	41.3	54.8	29.1	3.2	9.4
Ítem 3	17.7	17.3	21	42.3	58.1	29.6	3.2	10.8
Ítem 4	17.7	19.3	35.5	39.4	38.7	31.9	8.1	9.4

se refieran a la asociación, independencia y otros conceptos y propiedades. Observamos que se repiten los descritos por Estepa (1993), especialmente los que denomina *concepción local y causal de la asociación*. Nosotros no hemos podido ver la estabilidad de estas supuestas concepciones, por eso usamos el término *conflicto* que nos parece menos fuerte que el de concepción, pues en principio sólo implica una confusión por parte del alumno, más que una idea asumida con convicción. También aparecen nuevos conflictos no descritos por este autor, que clasificamos a continuación.

#### *Confusión entre conceptos o propiedades*

- Confusión de probabilidades con casos favorables. Se produce al calcular o comparar probabilidades. Puede ser originado por el lenguaje, pues a veces se usan como sinónimos posibilidad y probabilidad en el lenguaje ordinario, aunque no en matemáticas. Este conflicto también apareció en el trabajo de Díaz y De la Fuente (2005).
- Pensar que se pueden comparar muestras de distinto tamaño con base en las frecuencias absolutas. Este conflicto también se describe en Estepa (1993) y Arteaga (2008), quienes indican que se debe a una incomprensión del concepto de distribución, pues los estudiantes tratan de comparar distribuciones usando sólo frecuencias absolutas.
- Pensar que no se pueden comparar muestras de tamaños diferentes, error descrito por Watson (2006).
- Confusión del denominador: el cálculo riesgo relativo. Supone una falta de comprensión del concepto riesgo relativo y se produce cuando se comparan riesgos relativos, uno correspondiente a una distribución condicional por filas y otro por columnas, cuando se deberían comparar dos de la misma

Tabla 5. Conflictos semióticos en las diferentes estrategias

Tipo	Conflictos	Estrategia
Confusión de conceptos o propiedades	Confundir probabilidades con casos favorables	<i>SP1, SP2, SI5, SI6, SI8</i>
	Pensar que se pueden comparar muestras de distinto tamaño con base en las frecuencias absolutas	<i>SP1, SP2, SP3, SI1, SI2, SI6</i>
	Pensar que no se pueden comparar muestras de distinto tamaño	<i>SI4</i>
	Confundir el denominador en el cálculo riesgo relativo	<i>SP5</i>
	Confundir frecuencia con valor de la variable	<i>SI3, SI5</i>
	Confundir frecuencias conjuntas y condicionales	<i>SI3</i>
	Calcular la media de las frecuencias	<i>SI3, SI5</i>
	Confundir la escala de medida de la variable	<i>SI5</i>
	Confundir una tabla de contingencia con una tabla de frecuencias	<i>SI9</i>
Propiedades incorrectas de la asociación	Suponer que hay asociación si las frecuencias absolutas condicionales son distintas	<i>SP1, SI2</i>
	Confundir causalidad con asociación (concepción causal)	<i>SP2</i>
	Deducir la asociación sólo de una parte de los datos (concepción local)	<i>SI1, SI2, SI3, SI6, SI8</i>
	Pensar que la celda $d$ es contraria a la asociación	<i>SI5</i>
	Basar la asociación en la diferencia de frecuencias en las celdas situadas en las diagonales	<i>SI10</i>
	Basar la asociación sólo en las teorías previas	<i>SI11</i>
Propiedades incorrectas de la independencia	Suponer la equiprobabilidad en las celdas dobles en caso de independencia	<i>SP4, SI3</i>
	Suponer que la asociación implica igualdad de las celdas $a$ y $d$	<i>SI7</i>

distribución. No hemos encontrado este conflicto descrito en investigaciones previas.

- Confundir frecuencia con valor de la variable. Esta confusión indica incomprensión del concepto de distribución y fue descrita por Ruiz (2006).
- Calcular la media de las frecuencias. El estudiante supone que se puede calcular la media de las frecuencias, no comprendiendo el significado de la media, que no tiene sentido en variables cualitativas. Implica, por tanto, el conflicto siguiente. Este conflicto fue descrito por Mayén, Batanero y Díaz (2009) en su trabajo sobre la mediana.
- Confundir la escala de medida de la variable, suponiendo que la variable es cuantitativa en vez de nominal. Implica incomprensión de la diferencia entre estos tipos de escalas y variables. Este conflicto también fue descrito por Mayén, Batanero y Díaz (2009).
- Confundir una tabla de contingencia con una tabla de frecuencias. Los estudiantes no comprenden la finalidad de la tabla y la tratan como una de una sola variable. Por tanto, están confundiendo variable doble con variable simple y frecuencia conjunta con frecuencia marginal. No lo hemos encontrado descrito.

#### *Propiedades incorrectas atribuidas a la asociación*

- Suponer que hay asociación si las frecuencias absolutas condicionales son distintas. Hay asociación si las frecuencias relativas condicionales son diferentes. Los estudiantes generalizan incorrectamente una propiedad de las frecuencias relativas a las absolutas mediante un proceso incorrecto de generalización. No lo hemos encontrado descrito.
- Confundir causalidad con asociación. Este conflicto es detectado por Estepa (1993), quien lo denomina concepción causal de la asociación. Como hemos señalado, si existe causalidad, existe asociación, pero no al contrario. Los estudiantes confunden la doble implicación con la implicación simple.
- La asociación se deduce sólo de una parte de los datos. Supone la existencia de la concepción local sobre la asociación descrita por Estepa (1993). Aparece en todas las estrategias basadas en la comparación de sólo una parte de los datos.
- Pensar que la celda  $d$  es contraria a la asociación, no detectando el papel simétrico que tienen las celdas  $a$  y  $d$  para el estudio de la asociación. Ya Inhelder y Piaget (1955) describieron esta dificultad, indicando que estos sujetos están en una etapa primera de la comprensión de la asociación.

- Basar la asociación en la diferencia de frecuencias en las celdas situadas en las diagonales. En realidad, este conflicto está relacionado con el anterior, pues el sujeto no concibe el papel simétrico de las celdas a y d y de las celdas b y c para el estudio de la asociación.
- Basar la asociación en las teorías previas, mostrando el fenómeno de correlación ilusoria descrito en varias investigaciones (Jennings, Amabile y Ross, 1982; Alloy y Tabacnick, 1984; Wrigth y Murphy, 1984).

#### *Propiedades incorrectas atribuidas a la independencia*

- Suponer la equiprobabilidad en las celdas dobles en caso de independencia. Supone que, en caso de independencia, la frecuencia ha de ser igual en todas las celdas de la tabla. Por un lado, esto implica el sesgo de equiprobabilidad, descrito por varios autores. Por otro, una confusión entre las ideas de independencia y equiprobabilidad. No lo hemos encontrado descrito.
- Suponer que la asociación implica igualdad de las celdas a y d. Se comparan estas celdas suponiendo que han de ser iguales en caso de independencia. No lo hemos encontrado descrito.

## CONCLUSIONES

El estudio empírico descrito en este artículo muestra unos pobres resultados en las estrategias para realizar juicios de asociación en los estudiantes que ingresan en Psicología. Aunque la muestra es de tamaño moderado, los participantes fueron un grupo completo de los dos que integran los estudiantes de primer año en la Universidad de Huelva, cuyas características no son muy diferentes de las de los estudiantes de psicología en otras universidades españolas en cuanto a nota de acceso a los estudios y bachillerato cursado.

Al comparar con el estudio de Estepa (1993), realizado con estudiantes del Curso de Orientación Universitaria, se muestra en nuestro caso un menor porcentaje de estrategias parcialmente correctas para los cuatro ítems del cuestionario (tres tablas  $2 \times 2$ : independencia, asociación inversa, asociación directa; y una tabla  $2 \times 3$ : asociación directa). Las tres estrategias más frecuentes en ambos estudios fueron incorrectas, siendo en los dos casos la más frecuente “utilizar una única distribución condicional”.

Por otro lado, el análisis de los conflictos semióticos realizado en este estudio confirma la presencia de las concepciones local y causal de la asociación descritas



por Estepa en una proporción importante de estudiantes, quienes también en ocasiones se guían por sus teorías previas, mostrando la correlación ilusoria. El análisis realizado de conflictos semióticos amplía la variedad de dificultades descritas por Estepa y, al mismo tiempo, da una explicación menos exigente de estas dificultades, pues el conflicto semiótico supone un error de interpretación y no tiene el carácter de estabilidad y resistencia al cambio descrito en las concepciones.

En el trabajo encontramos algunas nuevas estrategias no descritas en la investigación previa. Además, se ofrece una clasificación de los conflictos detectados en los estudiantes en tres grupos, dependiendo de si se relacionan con la confusión de conceptos o respecto a la atribución de propiedades incorrectas a la asociación e independencia.

Pensamos que el análisis de dichos conflictos de los estudiantes es un primer paso para organizar acciones formativas adecuadas que permitan hacer aflorar dichos conflictos con la finalidad de superarlos.

## AGRADECIMIENTOS

Proyecto EDU2010-14947, becas FPU-AP2009-2807 (MCINN-FEDER), FPI BES-2008-003573 y FPU-AP2007-03222 (MEC-FEDER) y grupo FQM126 (Junta de Andalucía).

## REFERENCIAS BIBLIOGRÁFICAS

- Allan, L. G. y H. M. Jenkins (1983), "The effect of representations of binary variables on judgment of influence", *Learning and Motivation*, núm. 14, pp. 381-405.
- Alloy, L. B. y N. Tabachnik (1984), "Assessment of covariation by humans and animals: The joint influence of prior expectations and current situational information", *Psychological Review*, núm. 91, pp. 112-149.
- Arteaga, P. (2008), *Análisis de gráficos estadísticos elaborados en un proyecto de análisis de datos*, Trabajo fin de Master, Departamento de Didáctica de la Matemática.
- Barbancho, A. G. (1973), *Estadística elemental moderna*, Barcelona, Ariel (4a. ed., reimpresión de 1975).
- Bertin, J. (1967), *Semiologie graphique*, París, Gauthier-Villars.
- Chapman, L. J. y J. P. Chapman (1969), "Illusory correlation as an obstacle to the

- use of valid Psychodiagnostic signs”, *Journal of Abnormal Psychology*, núm. 74, pp. 271-280.
- Díaz, C. y E. I. de la Fuente (2005), “Conflictos semióticos en el cálculo de probabilidades a partir de tablas de doble entrada”, *Biaix*, núm. 24, pp. 85, 91.
- Díaz, J. y B. Gallego (2006), “Algunas medidas de utilidad en el diagnóstico”, *Revista Cubana de Medicina General Integrada*, vol. 22, núm. 1.
- Eco, U. (1977), *Tratado de semiótica general*, Barcelona, Lumen.
- Estepa, A. (1993), *Concepciones iniciales sobre la asociación estadística y su evolución como consecuencia de una enseñanza basada en el uso de ordenadores*, Tesis Doctoral, Universidad de Granada.
- Estepa, A. y C. Batanero (1995), “Concepciones iniciales sobre la asociación estadística”, *Enseñanza de las Ciencias*, vol. 13, núm. 2, pp. 155-170.
- Falk, R. (1986), “Conditional probabilities: Insights and difficulties”, en R. Davidson y J. Swift (eds.), *Proceedings of the Second International Conference on Teaching Statistics*, Victoria, Canada, International Statistical Institute, pp. 292-297.
- Font, V., J. D. Godino y B. D’Amore (2007), “An ontosemiotic approach to representations in mathematics education”, *For the Learning of Mathematics*, vol. 27, núm. 2, pp. 2-7.
- Inhelder, B. y J. Piaget (1955), *De la logique de l’enfant à la logique de l’adolescent*, París, Presses Universitaires de France.
- Jenkins, H. M. y W. C. Ward (1965), “Judgment of the contingency between responses and outcomes”, *Psychological Monographs*, núm. 79, pp. 1-17.
- Jennings, D. L., T. M. Amabile y L. Ross (1982), *Informal covariation assessment: Data-based versus theory-based judgments*, en D. Kahneman, P. Slovic y A. Tversky (eds.), *Judgment under uncertainty: Heuristics and biases*, Nueva York, Cambridge University Press, pp. 211-230.
- Mayén, S., C. Batanero y C. Díaz (2009), “Conflictos semióticos de estudiantes mexicanos en un problema de comparación de datos ordinales”, *Revista Latinoamericana de Investigación en Matemática Educativa*, vol. 12, núm. 2, pp. 151-178.
- Pérez Echeverría, M. P. (1990), *Psicología del razonamiento probabilístico*, Madrid, Ediciones de la Universidad Autónoma de Madrid.
- Ruiz, B. (2006), *Un acercamiento cognitivo y epistemológico a la didáctica del concepto de variable aleatoria*, Tesis de Maestría, CICATA, México.
- Schild, M. (2006), “Statistical literacy survey analysis: reading graphs and tables of rates percentages”, en B. Phillips (ed.), *Proceedings of the Sixth International Conference on Teaching Statistics*, Ciudad del Cabo, International Statistical

- Institute and International Association for Statistical Education. Online: [www.statauckland.ac.nz/~iase](http://www.statauckland.ac.nz/~iase).
- Shaklee, H. y M. Mins (1982), "Sources of error in judging event covariations: Effects of memory demands", *Journal of Experimental Psychology Learning, Memory and Cognition*, vol. 8, núm. 3, pp. 208-224.
- Shaklee, H. y D. Tucker (1980), "A rule analysis judgments of covariation between events", *Memory and Cognition*, núm. 8, pp. 459-467.
- Smedlund, J. (1963), "The concept of correlation in adults", *Scandinavian Journal of Psychology*, núm. 4, pp. 165-174.
- Watson, J. M. (2006), *Statistical literacy at school: Growth and goals*, Mahwah, NJ, Lawrence Erlbaum Associates.
- Wright, J. C. y G. L. Murphy (1984), "The utility of theories in intuitive statistics: The robustness of theory-based judgments", *Journal of Experimental Psychology General*, vol. 113, núm. 2, pp. 301-322.

## DATOS DE LOS AUTORES

### **Gustavo Cañadas**

Departamento de Didáctica de la Matemática,  
Universidad de Granada, Granada, España  
[gfuente@correo.ugr.es](mailto:gfuente@correo.ugr.es)

### **Carmen Batanero**

Departamento de Didáctica de la Matemática,  
Universidad de Granada, Granada, España  
[batanero@ugr.es](mailto:batanero@ugr.es)

### **José Miguel Contreras**

Departamento de Didáctica de la Matemática,  
Universidad de Granada, Granada, España  
[jmcontreras@ugr.es](mailto:jmcontreras@ugr.es)

### **Pedro Arteaga**

Departamento de Didáctica de la Matemática,  
Universidad de Granada, Granada, España  
[parteaga@ugr.es](mailto:parteaga@ugr.es)

