

# **Análisis de varianza elemental versus contraste de hipótesis: comprensión de las hipótesis estadísticas mediante la identificación y comparación de conflictos semióticos**

## **Elementary variance analysis versus hypothesis testing: understanding statistical hypotheses by identifying and comparing semiotic conflicts**

Osmar Vera

Universidad Nacional de Quilmes

### **Resumen**

La finalidad del trabajo es evaluar en forma comparada la manera en que los estudiantes establecen las hipótesis estadísticas en un problema de análisis de varianza y otro de prueba de hipótesis. Realizamos un estudio semiótico de las respuestas dadas por una muestra de 224 estudiantes universitarios de Psicología de la Universidad de Huelva, España a dos problemas abiertos, analizamos su capacidad de argumentación y las estrategias utilizadas al resolverlos. Propusimos como objetivo identificar y analizar comparadamente los conflictos semióticos al establecer las hipótesis estadísticas. Encontramos que comparten conflictos al usar la media muestral como parámetro, al confundir el campo de problemas y en la variable usada en el contraste. El marco teórico utilizado es el Enfoque Ontosemiótico de la cognición matemática, que viene desarrollando Godino y colaboradores desde el año 1994.

**Palabras clave:** análisis de varianza, comparación de conflictos semióticos, hipótesis estadísticas, estudiantes universitarios.

### **Abstract**

The aim of this work is performing a comparative analysis of the way in which the students establish the statistical hypotheses in a variance analysis problem and another hypothesis test problem. We analyse the students' argumentative capacity and the strategies used to solve the problems. To achieve this goal, we performed a semiotic study of the answers given by a sample of 224 university students of Psychology at the University of Huelva, Spain to two open problems. We aimed to identify and compare the semiotic conflicts involved in establishing statistical hypotheses. We find that the students share conflicts when using the sample mean as a parameter, in confusing the problem field and in the variable used in the test. The theoretical framework is the Ontosemiotic Approach to Mathematical Cognition developed by Godino and collaborators since 1994.

**Keywords:** Variance Analysis, comparison of semiotic conflicts, statistical hypothesis, university students.

## **1. Introducción**

El uso e interpretación de la inferencia estadística tiene un papel destacado en las investigaciones que involucran las ciencias humanas, incluyendo sociología, educación y psicología. En diversas revisiones (ver por ejemplo, Batanero y Díaz, 2006), se muestra que su utilización de la inferencia en estas investigaciones muchas veces no es adecuada. A pesar que el software hoy en día es facilitador de los cálculos estadísticos para la toma de decisión en la prueba, los estudiantes manifiestan dificultades en la comprensión del proceso, lo que los lleva finalmente a caer en decisiones erróneas (Harradine, Batanero y Rossman, 2011). Son muchas las investigaciones que describen errores en la comprensión de la inferencia en estudiantes universitarios (Cañadas,

---

Vera, O. (2017). Análisis de varianza elemental versus contraste de hipótesis: comprensión de las hipótesis estadísticas mediante la identificación y análisis de conflictos semióticos comparados. En J. M. Contreras, P. Arteaga, G. R. Cañadas, M. M. Gea, B. Giacomone y M. M. López-Martín (Eds.), *Actas del Segundo Congreso Internacional Virtual sobre el Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos*. Disponible en, [enfoqueontosemiotico.ugr.es/civeos.html](http://enfoqueontosemiotico.ugr.es/civeos.html)

Batanero, Díaz y Roa, 2012; Castro Sotos, Vanhoof, Van den Norrgate y Onghena, 2007; Vallecillos, 1994; Vera, Batanero, Díaz y López-Martín, 2016; Vera, Díaz y Batanero, 2011).

Consideramos que el correcto planteamiento tanto de la hipótesis nula como la alternativa es esencial para realizar un contraste, puesto que fallando en ello puede resultar la elección de un contraste no adecuado, una mala construcción de la zona crítica y, la toma de decisión final errónea (Vera et al., 2011; 2016).

Para este trabajo, se pasó una prueba de evaluación a los estudiantes de una muestra que más abajo se describe -con algunos problemas abiertos-, con la finalidad de evaluar su capacidad de argumentación y las estrategias utilizadas al resolver los problemas. Se realizó un estudio semiótico de esas respuestas en dichos problemas abiertos; nos apoyamos en nociones teóricas del enfoque Ontosemiótico de la cognición matemática (Godino, 2002; Godino, Batanero y Font, 2007).

Godino, Font y Wilhelmi (2008) destacan la existencia de cinco niveles del análisis didáctico, sobre el segundo nos centraremos en este trabajo. Este nivel de análisis se centra en los objetos y procesos que intervienen en la realización de las prácticas matemáticas y tiene como objetivo describir su complejidad onto-semiótica como un factor explicativo de los conflictos semióticos que en su realización se producen.

Seguimos el método usado en investigaciones previas, por ejemplo, Cañadas (2012) y Gea (2014). De este modo, comparamos la comprensión lograda por los estudiantes en el significado para el planteamiento de las hipótesis estadísticas en una prueba de hipótesis (P1) y del análisis de la varianza (P2). En lo que sigue describimos los fundamentos del trabajo en relación a la comprensión de las hipótesis en un contraste estadístico y el marco teórico utilizado; la metodología empleada, la discusión de los resultados y las conclusiones.

## **2. Fundamentos del trabajo**

### **2.1. Comprensión de las hipótesis en un contraste estadístico**

Chow (1996) señala como error más frecuente en la interpretación de las hipótesis estadísticas el confundirla con la hipótesis de investigación. A qué se refiere una y otra? Mientras que la hipótesis de investigación suele ser amplia y describe un rasgo inobservable (por ejemplo, que el rendimiento de los estudiantes es diferente en dos grupos), la hipótesis estadística, hace referencia a un parámetro de una distribución de probabilidad en una población de sujetos. A pesar de esta diferencia clara, Batanero (2000) indica que muchos investigadores extrapolan un resultado significativo directamente a la hipótesis de investigación. Los estudiantes también confunden las hipótesis nula y alternativa (Vallecillos, 1994). La autora encontró un 13% aproximadamente entre 436 estudiantes que confundieron la hipótesis nula con la alternativa al plantear las hipótesis en problemas de contraste de hipótesis sencillos. Otro 20% de estudiantes en esa investigación no discriminaron si la hipótesis del contraste se refiere al parámetro de la población o bien al estadístico muestral y plantearon sus hipótesis usando los estadísticos muestrales. Mientras que el parámetro de la población (por ejemplo la media de una distribución normal) es un valor constante y desconocido, el correspondiente estadístico (en el ejemplo, la media de una muestra tomada de dicha población) es conocido (pues obtenida la muestra, se puede calcular)

pero variable (ya que diferentes medias tomadas de la misma población pueden dar lugar a valores ligeramente distintos de las medias muestrales).

Vallecillos en su trabajo también propuso un problema abierto a los estudiantes, encontrando que sólo el 26% de los participantes planteó correctamente las hipótesis. Por otro lado, entrevistó a siete de ellos considerados brillantes por sus profesores, indicando que todos concebían el concepto de hipótesis como una afirmación sujeta a confirmación. Comprendían que las hipótesis nula y alternativa son complementarias pero tienen diferente papel en el contraste. Finalmente, aunque todos admitieron en la entrevista que la hipótesis se debe referir a un parámetro, algunos de ellos plantearon la hipótesis en función del estadístico muestral. Errores similares se describen asimismo en las investigaciones de Cañadas, Batanero, Díaz y Roa (2012), Castro Sotos et al. (2007) y en Vallecillos y Batanero (1997). En nuestro trabajo también encontramos que los estudiantes no consideran adecuadamente las poblaciones que se estudian, es decir consideran dos cuando en realidad se trata de una sola (P1), o menos de tres como en el caso de P2.

Interesados por estos resultados y por el hecho de que la comprensión y correcto planteamiento de las hipótesis en un contraste es un requisito necesario para finalizar e interpretar correctamente todo el procedimiento, el objetivo de este trabajo es ahondar un poco más en la comprensión de este tema por parte de los estudiantes. Más concretamente, y utilizando el marco teórico que se describe a continuación, queremos analizar comparativamente y con mayor profundidad los posibles planteamientos incorrectos de las hipótesis en un problema tanto de contraste de hipótesis como de análisis de varianza, ambos elementales.

En dos trabajos previos (Vera, Díaz y Batanero, 2011 y Vera, Batanero, Díaz y López-Martín, 2016) analizamos la formulación de hipótesis estadísticas en un contraste sobre la media de una población y en un modelo de análisis de varianza de medidas repetidas para tres poblaciones por estudiantes de Psicología, ambos apoyados en el mismo marco teórico, con la finalidad de identificar los conflictos semióticos relacionados con los objetos que intervienen en el planteamiento de las hipótesis estadísticas. Entre los conflictos que comparten los dos problemas, los que hemos identificado fueron: la utilización de la media muestral en lugar del parámetro, la confusión en el campo de problemas, y el no identificar un parámetro adecuado en ambas hipótesis, confundiendo variables dependiente con independiente.

## 2.2. Marco Teórico

En este trabajo nos basamos en ideas teóricas propuestas en el enfoque ontosemiótico (Godino, 2002; Godino, Batanero y Font, 2007) que concibe el significado de los objetos matemáticos o estadísticos (por ejemplo, los conceptos de hipótesis nula y alternativa) como un sistema complejo de prácticas, donde intervienen los siguientes tipos de objetos matemáticos: *Situaciones-problemas*, de donde surge el objeto: el concepto de hipótesis estadística surgirá de problemas de comparación de los parámetros una, de dos o más poblaciones. *Lenguaje*: términos, expresiones, notaciones, gráficos que se usan en el trabajo matemático (por ejemplo los símbolos usados para denotar los parámetros  $\mu$ ,  $\sigma$ , o los usados para hipótesis nula  $H_0$  y alternativa  $H_1$ ). *Conceptos*: por ejemplo, población y muestra, estadístico y parámetro, región crítica y de aceptación, factor, niveles del factor, interacción. *Propiedades*: por ejemplo, que las hipótesis nula y alternativa son complementarias o que las hipótesis se formulan en

función del parámetro asociado a alguna distribución. *Procedimientos*; como los requeridos para construir las regiones críticas y de aceptación. *Argumentos* y *razonamientos*: usados para justificar o explicar a otra persona las proposiciones y procedimientos.

Para un objeto matemático (en este caso la hipótesis estadística) en este marco teórico se distingue entre significado institucional y personal. El significado institucional incluye las prácticas matemáticas que se comparten en una institución, mientras que el significado personal estaría formado por las prácticas personales de un sujeto, alguna de las cuáles podrían no coincidir con las pretendidas en la institución.

Godino (2002) señala que, en las prácticas matemáticas, intervienen objetos ostensivos (símbolos, gráficos, etc.) y no ostensivos (que evocamos al hacer matemáticas); los símbolos (significantes) remiten a entidades conceptuales (significados). Estas representaciones tienen mucha importancia para facilitar la enseñanza y el aprendizaje, pero a veces causan dificultades en los estudiantes. Godino, Batanero y Font (2007) toman de Eco (1995) la noción de función semiótica como una "correspondencia entre conjuntos", donde intervienen tres componentes: *un plano de expresión* (objeto inicial, considerado frecuentemente como el signo); *un plano de contenido* (objeto final, considerado como el significado del signo, esto es, lo representado, lo que se quiere decir, a lo que se refiere un interlocutor) y *un criterio o regla de correspondencia* (esto es un código interpretativo que relaciona los planos de expresión y contenido).

Esta idea de función semiótica sirve para explicar algunas dificultades y errores de los estudiantes. Los autores denominan conflicto semiótico a las interpretaciones de expresiones matemáticas por parte de los estudiantes (significados personales) que no concuerdan con las que el profesor trata de transmitir (significados institucionales). Los autores indican que estos conflictos semióticos producen errores en los estudiantes, que no son debidos a falta de conocimiento, sino a una interpretación incorrecta de expresiones matemáticas. Utilizaremos este marco para identificar conflictos que comparten las respuestas para ambos problemas propuestos.

### 3. Método

La muestra estuvo formada por 224 estudiantes de segundo año de Psicología de la Universidad de Huelva, España (con edades de 19 y 20 años). Con la finalidad de evaluar la comprensión del estudiante de los conceptos de hipótesis nula y alternativa y su competencia para identificar las hipótesis adecuadas en el análisis de varianza y prueba de hipótesis propiamente dicha en forma comparada, se les propuso la tarea que se presenta en la Figura 1, que los estudiantes respondieron por escrito e individualmente.

Los datos fueron recogidos como parte de una evaluación en una asignatura de análisis de datos que cursaron el segundo año, en la que estudiaron muestreo, estimación de intervalos de confianza, contraste de hipótesis sobre medias y proporciones, y análisis de varianza. Los estudiantes habían cursado el primer año estadística descriptiva y probabilidad.

Para el proceso de elaboración de los problemas se siguieron una serie de recomendaciones psicométricas habituales. En cada problema se categorizan las respuestas en correctas, parcialmente correctas e incorrectas, para mostrar la variedad de las mismas. Esta categorización nos permite analizar las estrategias y argumentos de los

estudiantes, evaluando su comprensión. Tomamos las respuestas incorrectas para ambos problemas y comparamos los conflictos semióticos de ambas a fin de evaluar sus similitudes.

En **P1**, de la Figura 1, se presentó un problema sobre contraste de hipótesis propiamente dicho, donde se plantea una situación práctica en relación con la velocidad lectora de palabras por minuto que son capaces de leer un grupo de estudiantes. Siendo que la investigación propone algunos parámetros, un profesor quiere especificar cómo se relacionan estos con su grupo. Recoge medidas de velocidad lectora de estudiantes, sometiendo los resultados al contraste. Se sabe que los niños de esa edad tienen una velocidad lectora de 40 palabras por minuto, quiere saber si sus datos aportan suficiente evidencia para rechazar o no ese rendimiento para sus estudiantes, de donde la hipótesis nula es:  $H_0: \mu = 40$ . La hipótesis alternativa es su complemento, entonces esta es:  $H_1: \mu \neq 40$ .

**P1.** Se sabe por diversos trabajos de investigación que los niños de seis años tienen una velocidad lectora media de 40 palabras por minuto, con varianza igual a 16. Un profesor quiere saber si los niños de su clase se sitúan o no en la media de palabras por minuto. Para ello mide la velocidad de lectura en los 25 niños de su clase, obteniendo una media de 43 palabras por minuto.

**P2.** Se ha llevado a cabo una investigación para estudiar si las técnicas de reducción de estrés tienen efecto sobre la ansiedad precompetitiva en atletas. Para ello se seleccionaron 5 atletas con alta ansiedad y les enseñaron estas técnicas. Se tomaron tres medidas de ansiedad: A1: antes de la enseñanza; A2: durante la enseñanza; A3: después de la enseñanza.

Define las hipótesis estadísticas que se someten a contraste tanto para **P1** como para **P2**

Figura 1. Enunciado de problemas abiertos presentados

Para **P2**, de la Figura 1, se recogen medidas de ansiedad de cinco atletas para tres momentos distintos: antes, durante y después de la aplicación de ciertas técnicas, y se desea saber el efecto de este factor con tres niveles (momento en que se realiza la medición) sobre las respuestas. Puesto que la hipótesis nula en un contraste de hipótesis es siempre la hipótesis de no efecto, en este problema la hipótesis nula debe indicar que la respuesta media para los tres momentos de los sujetos no varía, y por tanto, las técnicas de reducción del estrés no producen efecto sobre la ansiedad. La hipótesis nula adecuada posible sería:  $H_0: \mu_{antes} = \mu_{durante} = \mu_{después}$

La hipótesis alternativa es la que le interesa al investigador poner a prueba, en este caso, que las técnicas tienen un efecto sobre la ansiedad, o lo que es lo mismo, si se produce una diferencia en la ansiedad media en algún momento de la medición (antes-durante-después). Puesto que la hipótesis alternativa es el complemento de la nula, se debe expresar en forma general, sin indicar cuál de las tres medias es diferente de las otras. Una posible forma de hacerlo sería la siguiente:

$H_1 \equiv$  al menos una de las medidas de ansiedad en las tres momentos de medición es diferente de las otras

#### 4. Resultados y discusión

Recogidos los datos, se llevó a cabo un análisis cualitativo en que, mediante un proceso cíclico e inductivo se compararon las respuestas semejantes entre sí en cada uno de los

problemas por separado, para llegar a una categorización. El proceso se repitió varias veces, y fue revisado por el autor del trabajo, con ayuda de compañeros del grupo de investigación, discutiendo los casos de desacuerdo para mejorar la fiabilidad del proceso. Luego se compararon las categorías de respuestas incorrectas entre ambos problemas, analizando las similitudes en los errores cometidos.

Se clasificaron las respuestas como correctas, parcialmente correctas e incorrectas. En **P1** se encontraron diez categorías para las respuestas incorrectas, seis están dentro de aquellas en que usan el parámetro poblacional para contrastar y otras cuatro en las que usan la media muestral como parámetro. En **P2** las respuestas incorrectas se clasificaron en seis categorías distintas, teniendo en cuenta los estudiantes que consideraron solo una ó dos poblaciones, los que no consideran el momento en que se mide la ansiedad, o si se establecen las hipótesis en función de los estadísticos muestrales.

En lo que sigue (Tabla 1, 2 y 3) presentamos los conflictos semióticos de las categorías comparadas para ambos problemas que guardan similitudes en sus errores, ya sea por usar la media muestral como parámetro, que involucren más o menos poblaciones que las presentes en el problema y cuando confunden la variable de comparación. Mostramos y comentamos un ejemplo en cada una de ellas, realizando su análisis semiótico. Finalmente presentamos y discutimos los resultados, para finalizar con las conclusiones.

Tabla 1. Confusión entre media poblacional y muestral.

Expresión	Contenido
<b>P1</b> $H_1: \bar{X}_1 \neq \bar{X}_2$ $H_0: \bar{X}_1 = \bar{X}_2$	<ul style="list-style-type: none"> <li>- El alumno realiza una interpretación incorrecta del enunciado (P1), asumiendo que existen dos poblaciones, mientras que lo hace correctamente para P2, son tres las poblaciones (conceptos).</li> <li>- Escribe sus hipótesis utilizando una notación adecuada, tanto para hipótesis nula, la igualdad como la alternativa, desigualdad (concepto y representación).</li> </ul>
<b>P2</b> $H_0: \bar{X}_A = \bar{X}_B = \bar{X}_C$ $H_1: \bar{X}_A \neq \bar{X}_B \neq \bar{X}_C$	<ul style="list-style-type: none"> <li>- <i>Conflicto</i> al plantear un contraste de diferencia entre dos poblaciones (P1) (confusión del campo de problemas).</li> <li>- Discrimina la hipótesis nula como puntual y la hipótesis alternativa como aquella que quiere probar (propiedades).</li> <li>- Genera <i>conflicto</i> al plantear las hipótesis en término de un estadístico y no de un parámetro (concepto). Un posible <i>conflicto</i> también es la confusión entre población y muestra (conceptos).</li> <li>- Relacionado con el último anterior, presenta el <i>conflicto</i> consistente en confundir la media muestral que es un estadístico con un parámetro que es la media poblacional (conceptos).</li> </ul>

En primer lugar fue frecuente confundir el estadístico con el parámetro, estableciendo las hipótesis en términos del primero como se muestra en un ejemplo para cada problema en la Tabla 1. A modo de conclusión, los estudiantes de esta categoría confunden niveles de análisis que se corresponden con media poblacional y muestral, donde puede subyacer también la confusión entre población y muestra. Dicho error ha sido reportado en las investigaciones de Vallecillos (1994); Castro Sotos et al. (2007) y Cañadas et al. (2012).

Tabla 2. Confusión en el campo de problemas

Expresión	Contenido
<b>P1</b> $H_0 \equiv \mu_1 = \mu_2$ $H_1 \equiv \mu_1 \neq \mu_2$	<ul style="list-style-type: none"> <li>- El alumno realiza una interpretación incorrecta del enunciado (P1 y P2), asumiendo que existen dos poblaciones, mientras para P1 está presente una sola y están presentes tres para P2 (conceptos). Escribe sus hipótesis utilizando una notación adecuada (representación).</li> </ul>
<b>P2</b> <b>a.</b> $H_1: \text{Mont} > \text{N de gestiones}$ $H_0: \text{Mont} \leq \text{N de gestiones}$	<ul style="list-style-type: none"> <li>- Conflicto al plantear un contraste de diferencia entre dos poblaciones (P1 y P2a.) (confusión del campo de problemas).</li> <li>- Conflicto al plantear un análisis de varianza como un contraste de una media (P2b.) y de dos medias (P2a.) (confusión del campo de problemas).</li> </ul>
<b>b.</b> $H_0: \mu_{TIE} = 0$ $H_1: \mu_{TIE} \neq 0$	<ul style="list-style-type: none"> <li>- Discrimina la hipótesis nula como puntual (P1 y P2b.) y la hipótesis alternativa como aquella que quiere probar (propiedades).</li> <li>- En todos los casos utiliza un parámetro poblacional para la comparación (concepto y expresión).</li> <li>- Reconoce la igualdad para la hipótesis nula (P1 y P2), genera conflicto en (P2a.) puesto que no la considera puntual (propiedad).</li> <li>- Conflicto al considerar una prueba de cola superior en un análisis de varianza (P2a.) (campo de problemas).</li> </ul>

Los ejemplos analizados en la Tabla 2 corresponde a la confusión del tipo de contraste a realizar, Esto llevará a usar un estadístico inapropiado (por ejemplo, usar la prueba t en lugar del Análisis de la varianza).

Tabla 3. Confusión de la variable cuando identifican la hipótesis

Expresión	Contenido
<b>P1</b> $H_0: \mu = 6$ $H_1: \mu \neq 6$	<ul style="list-style-type: none"> <li>- El alumno interpreta incorrectamente el enunciado (conflicto) para definir las hipótesis (proceso de interpretación).</li> <li>- Reconoce la igualdad para la hipótesis nula, aunque en (P2) no la considera puntual (propiedad).</li> </ul>
<b>P2</b> $H_0 \equiv T_1 \leq A$ $H_1 \equiv T_1 > A$	<ul style="list-style-type: none"> <li>- Conflicto al confundir la variable bajo estudio (P1), plantea la comparación de variable dependiente con la independiente (P2) (particularización del concepto de variable, lenguaje y argumentación)</li> <li>- Conflicto, no logra identificar un parámetro adecuado para contrastar (P2) (lenguaje).</li> <li>- Compara dos momentos en el análisis (P2), no reconoce que debe usar algún modelo de análisis de varianza (campo de problemas)</li> </ul>

El ejemplo analizado en la Tabla 3 es específico del análisis de varianza y consiste en confundir las variables dependientes e independientes en el mismo. El análisis semiótico realizado de cada uno de estos ejemplos muestra la complejidad del establecimiento de las hipótesis y que, en general, los estudiantes tienen más de un conflicto, en caso de error en su formulación.

### *Conflictos semióticos comparados al establecer las hipótesis estadísticas*

En el análisis semiótico que se ha ido detallando anteriormente (Tablas 1,2 y 3), se han encontrado los siguientes conflictos semióticos (Godino, 2002), que también hemos cuantificado para conocer la medida de estudiantes que lo tienen:

- *Conflicto en la utilización del parámetro.* Un 20,4% de los estudiantes eligen la media muestral y no la poblacional, presentando un conflicto a la hora de determinar sobre que parámetro se contrasta. Este conflicto implica una confusión entre *estadístico* y *parámetro*, aunque reconocen la comparación de los tres momentos de análisis (P2), no así en (P1) donde suponen la existencia de dos muestras. Este conflicto estuvo presente también en un 12% en la investigación de Cañadas et al. (2012). El porcentaje de estudiantes que usan la media muestral en lugar del parámetro poblacional es similar al encontrado en la investigación de Vallecillos (1994) que es un 18,7%.
- *Conflicto en el campo de problemas* al generar un contraste para dos poblaciones (P1), o dos poblaciones y una población (P2) en Tabla 2. El mismo *conflicto* se repite al comparar las variables dependientes e independientes en un 16% de la muestra (P2) tanto en Tabla 2 como en Tabla 3. Este conflicto solo se ha encontrado en la investigación de Arteaga (2011) al comparar gráficos estadísticos, pero no en relación al contraste de hipótesis y es específico del análisis de varianza. Un 20,1% de la muestra tiene esta confusión.
- *Conflicto al generar una prueba de cola superior* (Tabla 2 P2a. y Tabla 3, en P2) para una análisis de varianza, suponemos confunde el campo de problemas con pruebas de hipótesis para contrastes unilaterales, tanta de una como de dos muestras (P2 en Tabla 2 y en Tabla 3).
- *Conflicto* al no lograr identificar un parámetro adecuado (P2 en Tabla 3), se trata de un 24,6% de la muestra. Compara la técnica para bajar el nivel de estrés con nivel de ansiedad, utilizando notación no relacionada con la que se reconoce en la institución.

## **5. Conclusiones**

En este trabajo no se han encontrado errores consistentes en formular hipótesis que no cubren el espacio paramétrico (7,7% en Vera et al, 2011), estimamos que los estudiantes incurren este error cuando se trata de problemas para contraste de hipótesis propiamente dichos. Tampoco se presentó coincidencia en el planteo de hipótesis usando una combinación lineal entre los parámetros (2,7% en Vera et al, 2016), pensamos que esta confusión la tienen los estudiantes a la hora de plantear modelos para el análisis de varianza. Se deberían estudiar cómo razonan los estudiantes en las pruebas a posteriori (luego de rechazar la hipótesis nula en un análisis de varianza) puesto que quizá esta confusión provenga de ese hecho.

Uno de los conflictos que se presenta en coincidencia para ambos problemas es la confusión entre estadístico y parámetro, que se manifiesta cuando los estudiantes plantean hipótesis usando la media muestral en lugar de la poblacional, es decir no reconocen que el estadístico es una variable aleatoria y no un parámetro para contrastar. Coincidimos con la investigación de Cañadas et al. (2012), Chow (1996) y Batanero (2000); Vallecillos (1994), también reportó errores en este sentido.



Suele ser muy difícil de comprender para los estudiantes el concepto de distribución muestral, puesto que es mucho más abstracto que los de distribución de una población o de una muestra (Harradine, Batanero y Rossman, 2011), consideramos que se debería insistir en la enseñanza de la inferencia en general. Siguiendo la sugerencia de estos autores, sería importante una intervención animando a los estudiantes a razonar sobre los comportamientos hipotéticos de más de una muestra tomadas de la misma población (usando poblaciones conocidas por los estudiantes). Estas conjeturas podrían contrastarse con actividades que involucren la toma repetida de muestras aleatorias, utilizando alguna de las muchas aplicaciones interactivas disponibles en la *web* junto a software libres muy potentes para estadística, tal como el R.

Otras coincidencias en conflictos que encontramos para ambos problemas están relacionados con la confusión en el campo de problemas y el uso de las variables (específicamente entre dependiente e independientes). No encontramos investigación previa que reporte estas coincidencias.

En resumen, el presente trabajo aporta nueva información para apoyar la necesidad de revisar la enseñanza de la inferencia estadística y más concretamente, en lo relativo al planteamiento de hipótesis estadísticas tanto en pruebas de hipótesis como en análisis de varianza, pues la escasa investigación en la enseñanza para este objeto indica que no se le ha prestado la debida atención. Entendemos que atender a estos conflictos es fundamental, puesto que provocarían dificultades a la hora de profundizar el estudio de estos objetos estadísticos y para la toma de decisiones, es decir el hecho de plantear una hipótesis usando la media muestral en lugar de la poblacional aportaría una confusión al calcular el estadístico de prueba, por ende podría no caer en la zona correcta y así provocar el error.

Sería entonces necesario ayudar a los estudiantes en la construcción del razonamiento inferencial, planteando actividades informales como se proponen en Batanero y Borovcnik (2016) antes de iniciar el aprendizaje formalizado de los contrastes de hipótesis. Creemos que nuestros resultados contribuirán a la mejora de la enseñanza del tema y ojalá incentiven a otros investigadores a proseguir analizando las dificultades de los estudiantes en la inferencia estadística en general.

## Referencias

- Batanero, C. (2000). Controversies around the role of statistical tests in experimental research. *Mathematical Thinking and Learning*, 2(1-2), 75-98.
- Batanero, C. y Borovcnik, M. (2016). *Statistics and probability in high school*. Rotterdam: Sense Publishers.
- Batanero, C. y Díaz, C. (2006). Methodological and didactical controversies around statistical inference. *Actes des 36 Journées de la Societe Francaise de Statistique*. Paris: Societé Française de Statistique [CD-ROM].
- 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., Díaz, C. y Roa, R. (2012). Psychology students' understanding of the chi-squared test. *Statistique et Enseignement*, 3 (1), 3-18.
- Castro Sotos, A. E., Vanhoof, S., Van den Nororgate y W. y Onghena, P. (2007). Student's misconceptions of statistical inference: A review of the empirical evidence form research on statistical education. *Educational Research Review*, 2 (2), 98-113.

- Chow, L. S. (1996). *Statistical significance: Rationale, validity and utility*. London: Sage.
- Eco, U. (1995). *Tratado de semiótica general*. Barcelona: Lumen. (Trabajo original publicado en 1967)
- Gea, M. M. (2014). *La correlación y regresión en Bachillerato: análisis de libros de texto y del conocimiento de los futuros profesores*. Tesis doctoral. Universidad de Granada.
- Godino, J. D. (2002). Un enfoque ontológico y semiótico de la cognición matemática. *Recherches en Didactique des Mathematiques*, 22 (2 y 3), 237-284.
- 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.
- Godino, J. D., Font, V. y Wilhelmi, M. R. (2008). Análisis didáctico de procesos de estudio matemático basado en el enfoque ontosemiótico. *Publicaciones*, 38, 25-48.
- Harradine, A., Batanero, C. y Rossman, A. (2011). Students and teachers' knowledge of sampling and inference. En C. Batanero, G. Burrill y C. Reading (Eds.), *Teaching statistics in school mathematics. Challenges for teaching and teacher education* (pp. 235-246). New York: Springer.
- Vallecillos, A. (1994). *Estudio teórico experimental de errores y concepciones sobre el contraste de hipótesis en estudiantes universitarios*. Tesis Doctoral. Universidad de Granada.
- Vallecillos, A. y Batanero, C. (1997). Conceptos activados en el contraste de hipótesis estadísticas y su comprensión por estudiantes universitarios. *Recherches en Didactique des Mathematiques*, 17, 29-48.
- Vera, O., Díaz, C. y Batanero, C. (2011). Dificultades en la formulación de hipótesis estadísticas por estudiantes de Psicología. *Unión*, 27, 41- 61.
- Vera, O., Batanero, C. y Díaz, C. y López-Martín, M. M. (2016). Assessing psychology students' difficulties in elementary variance analysis. *Diálogo Educativo*, 16(48), 487 – 511