

CUESTIONES METODOLÓGICAS EN LA INVESTIGACIÓN EDUCATIVA

ANGUSTIAS VALLECILLOS JIMÉNEZ

Universidad de Granada

QUINTO SIMPOSIO DE LA SOCIEDAD ESPAÑOLA DE
INVESTIGACIÓN EN EDUCACIÓN MATEMÁTICA
Almería, Septiembre 2001

CUESTIONES METODOLÓGICAS EN LA INVESTIGACIÓN EDUCATIVA



ANGUSTIAS VALLECILLOS JIMÉNEZ

Universidad de Granada

Sería quizás prematuro especular en este momento sobre la cuestión de si los métodos de la ciencia abstracta podrán en el futuro ser útiles en la investigación de los problemas sociales en una medida equivalente a como lo han sido en varios campos de la investigación física. Mas que poseemos teóricamente en todos los casos, y prácticamente en la medida en que la requerida labor del cálculo puede ser suministrada, los medios de desarrollar a partir de los datos estadísticos las semillas de verdades generales que yacen enterradas entre la masa de números, es una opinión que puede afirmarse con toda seguridad.

George Boole, *Investigación sobre las leyes del pensamiento*, pp. 25-26.

RESUMEN

En este trabajo pretendemos sintetizar algunas cuestiones de método aplicables a la investigación educativa. Para ello reflexionamos sobre el método seguido para la realización de una amplia investigación de referencia, Vallecillos (1994), que pertenece al campo de la educación estadística. Es un ejemplo de lo que podemos llamar ‘método estadístico’ que puede aplicarse como ‘modelo’ en la investigación educativa en general. Se incluyen también, a modo de ejemplo de su funcionamiento, los resultados obtenidos en esa investigación sobre la comprensión de un concepto clave en los contrastes de hipótesis como el nivel de significación.

ABSTRACT

In this paper we try to synthesize some methodological questions applicable to educational research. With this aim we reflect on the method followed in a wide reference research, Vallecillos (1994), that belongs to the field of statistical education. This research is an example of what we can call ‘statistical method’ that can be applied as ‘model’ in educational research in general. We also include, as an example of its functioning, the results obtained in that research about the understanding of a key concept in hypothesis testing as is the significance level.

INTRODUCCIÓN

El método científico en la investigación

Entendemos por ‘método’ cualquier conjunto de acciones ordenadas con las que pretendemos conseguir unos resultados y por ‘metodología’ la aplicación coherente de un método. La problemática metodológica viene de antiguo, fue planteada por Bacon con su defensa de la inducción y un referente

importante para nosotros es Descartes con su planteamiento de la duda metódica. En nuestro ámbito de la investigación es preciso citar también a Galileo que introdujo la medición como forma de observación exacta del mundo natural. Es de todos conocido que lo que hoy llamamos *método científico* es el método inspirado en la Física, ya que esta ciencia consiguió rápidos y excelentes resultados con su método y se convirtió así en ejemplo a seguir por todas las demás. En síntesis, este método consiste en observar los hechos significativos, sentar hipótesis que los expliquen satisfactoriamente, deducir de ellas consecuencias que puedan someterse a prueba por la observación para aceptar como provisionalmente verdadera aquella hipótesis que soporte mejor la prueba. El desarrollo de diversas ciencias a partir de la segunda mitad del siglo XVIII ha supuesto también el desarrollo de diversas metodologías específicas. La discusión sobre el método ha sido especialmente importante en las ciencias humanas por la complejidad de su objeto y por la coincidencia en ellas de sujeto y objeto del conocimiento lo que hace difícil o imposible la aplicación en ellas de los métodos de observación de otras ciencias; el hecho de que estas ciencias sean de desarrollo muy reciente hace que esta discusión esté todavía en gran parte abierta.

El método estadístico en la investigación en Didáctica de la Matemática

El campo de las matemáticas, como disciplina 'objetiva' ha sido siempre campo neutral para la observación de fenómenos educativos y eso ha hecho, por otra parte, que el área de la Didáctica de las Matemáticas sea un buen observatorio de fenómenos educativos e investigaciones sobre los procesos de enseñanza y aprendizaje.

La investigación educativa puede afectar a un número de problemas tan grande y a su vez tan sumamente diferentes unos de otros que pretender encontrar un esquema metodológico válido para todo tipo de propósito investigador se nos antoja una idea poco realista. Existen ya clasificaciones más o menos exhaustivas de los tipos de investigaciones que pueden realizarse, atendiendo al método a emplear, a su grado de aplicación, al tipo de datos disponibles, al tipo de variables analizadas, al objetivo al que se dirigen, etc., como puede verse, por ejemplo, en Bisquerra (1989).

Se nos antoja bastante probable que, en general, una vez bien definido el problema de investigación, ningún esquema metodológico conocido sea totalmente adecuado a la situación concreta que tenemos planteada. Este es el momento en el que el investigador debe adaptar o diseñar una metodología a la medida de las necesidades de su propia investigación. Para ello, es necesario determinar los diferentes elementos a tener en cuenta, cuáles de ellos son indispensables y decidir entre los diseños metodológicos conocidos que aspecto de cada uno de ellos es pertinente al efecto de conseguir los objetivos que se persiguen. En caso contrario el investigador deberá diseñar su propio esquema metodológico que, probablemente, será una mezcla de elementos de otros esquemas anteriores con alguna parte original, de manera que pueda responder finalmente de una manera más adecuada a los objetivos concretos de investigación definidos en el problema planteado. En este sentido, considero que el esquema metodológico que se describe aquí, es completo ya que emplea distintos tipos de muestras, variables, técnicas de obtención de datos, tipos de análisis de los mismos, etc., que es posible que no sea necesario utilizar íntegramente en todo tipo de investigaciones, sino que también será aplicable utilizando solamente algunas de sus partes.

Se nos antoja difícil hablar de la investigación en abstracto, así que en este trabajo sólo pretendemos sintetizar algunas cuestiones metodológicas en relación con la investigación educativa sobre la base de nuestra propia experiencia investigadora y en nuestro campo concreto de investigación.

Lo que sigue es simplemente un intento de sistematización terminológica con el fin de conseguir un esquema metodológico aplicable a la investigación educativa, ejemplificado en una primera investigación concreta y que se está empleando en la actualidad en otras investigaciones en curso a través de Proyectos de Investigación subvencionados por el Ministerio de Educación español. Aunque la investigación de

partida, que empleamos como base para 'teorizar' sobre ella, es un trabajo realizado en el campo concreto de la educación estadística, (Vallecillos, 1994), pretendemos abstraer los aspectos metodológicos de la misma con el fin de construir un esquema aplicable a la investigación en otros campos de la investigación educativa ya que podemos considerarla un ejemplo de *método estadístico*, de aplicación general. Se incluye, además, una parte de la investigación, la que se refiere al concepto de nivel de significación, para que se pueda valorar mejor la utilidad de la metodología seguida en orden a conseguir los objetivos previstos en el problema de investigación planteado.

DESCRIPCIÓN DE LA INVESTIGACIÓN DE REFERENCIA

Esta investigación, (Vallecillos, 1994), tuvo como meta realizar una valoración comprensiva y sistemática de las concepciones, dificultades y errores de los estudiantes universitarios en el aprendizaje del tema del contraste de hipótesis estadísticas. Se llevó a cabo entre los años 1990 y 1994 y participaron en ella un total de 436 estudiantes de catorce grupos y siete titulaciones distintas de la Universidad de Granada. La toma de datos se realizó mediante un cuestionario sobre contenidos conceptuales y procedimentales que consta de 20 ítems y un problema de aplicación. En seis de los ítems se le pide al alumno que razone su respuesta con el fin de valorar mejor su comprensión del enunciado y nuestra interpretación de su respuesta.

Posteriormente se realizó una entrevista semiestructurada a un grupo de siete estudiantes seleccionados intencionalmente. El objetivo general pretendido es el de proporcionarnos información complementaria con el fin de profundizar en nuestra comprensión sobre algunos errores 'básicos' detectados en el aprendizaje de los alumnos, así como tratar de identificar sus causas. También pretendemos asegurarnos de la comprensión y de la interpretación que hemos dado a algunas de las respuestas de los alumnos y de la estabilidad de sus respuestas.

ESQUEMA METODOLÓGICO: ELEMENTOS MÍNIMOS QUE COMPORTA UNA INVESTIGACIÓN EN EL CAMPO DE LA EDUCACIÓN

Un correcto enfoque metodológico de la investigación educativa comporta el análisis detallado de los siguientes elementos mínimos: objetivos que se persiguen, instrumento de recogida de datos y análisis de los mismos. A pesar de que pretendemos ser muy sintéticos, es necesario que profundicemos un poco más en los diversos aspectos que cada uno de estos elementos conllevan.

El cierre del proceso es la elaboración de un informe final de la investigación que recoja con todo detalle todos los aspectos de antecedentes, planteamiento y ejecución del proyecto, decisiones tomadas en el curso de su desarrollo, limitaciones de las conclusiones obtenidas así como líneas abiertas a la continuación del trabajo e implicaciones en otros ámbitos como, en nuestro caso, para la enseñanza del tema objeto de la investigación.

A) Problema de investigación, objetivos, hipótesis, estado de la cuestión

Determinación del problema de investigación

Aún a riesgo de que esta afirmación parezca una perogrullada, la definición del problema de investigación es la primera cuestión a resolver. Muchas veces nos encontramos con un problema no bien

definido que, por tanto, no podrá estar nunca bien resuelto. La total determinación y clarificación inicial del problema, nos llevará a una buena determinación del estado de la cuestión, hipótesis de partida, objetivos iniciales, etc., indispensables desde el punto de vista metodológico.

Esta es la fase de transformación de un interés inicial vago del investigador en un área problemática en un verdadero problema de investigación cuyos interrogantes estén perfectamente delimitados y concretados.

Estudio teórico previo

En nuestro trabajo el estudio teórico inicial se hizo con mucho detenimiento. Se estudió primero el aspecto conceptual del tema, los distintos enfoques en la teoría de los contrastes de hipótesis, Fisher y Neyman-Pearson, así como sus perspectivas histórica y epistemológica. Se estudio el problema de los contrastes también desde la perspectiva de la enseñanza analizando hasta seis elementos de significado, Godino y Batanero (1994), de síntesis de los conceptos implicados en ellos. También se hizo un minucioso análisis a priori de los contenidos incluidos en el estudio. Es especialmente destacable el análisis de contenido realizado por dos razones básicas: a) para delimitar claramente el problema de investigación a un contenido perfectamente concretado; b) para dotar de validez de contenido al instrumento utilizado en la toma de datos. Posteriormente, estos análisis teóricos sobre los distintos enfoques así como las cuestiones epistemológicas analizadas, nos han permitido entender mejor algunas de las respuestas de los estudiantes y 'explicar' en parte los múltiples problemas derivados del uso de las técnicas de contraste de hipótesis en la investigación experimental.

En general, creemos muy importante el estudio teórico previo del problema, que incluirá los aspectos de contenido científico, filosóficos, históricos y didácticos del tema. Además de situar el problema de investigación en su marco teórico y científico correspondiente, este tipo de estudios son indispensables para comprender las respuestas de los alumnos y profundizar en las dificultades de tipo epistemológico o didáctico, errores de aprendizaje, etc., presentes en todos los procesos de enseñanza y aprendizaje.

Antecedentes de la investigación

Las investigaciones previas sobre el tema son el punto de partida de las nuevas, de manera que un análisis crítico de las mismas, clasificación, conceptos estudiados, aspectos complementarios, caso de otros niveles o situaciones de enseñanza, etc., es un trabajo minucioso pero indispensable para emprender cualquier trabajo de investigación. También servirá para evitarnos descubrir el Mediterráneo, ya descubierto.

Objetivos, variables e hipótesis

Una vez planteado claramente el problema de investigación, determinamos los objetivos generales y concretos a conseguir.

Las variables a considerar, número y tipo, se deducen de todo lo anterior. Usualmente en la investigación educativa las variables a observar son de tipo cualitativo o cuantitativo y, en cada caso, se llevará a cabo un tipo de análisis adecuado al caso. Existen distintas clasificaciones de las variables actualmente según su forma de medida y es necesario estar permanentemente atentos a esta cuestión ya que podemos considerar que, a medida que una disciplina avanza la medida de las variables implicadas es progresivamente más sofisticada. Una clasificación de las variables y su medida puede verse en Afifi y Azen (1990).

El papel de las hipótesis en la investigación experimental y educativa es un importante elemento a tener en cuenta. En general, una hipótesis es una afirmación que suponemos cierta, esto es, una afirma-

ción que expresa nuestra interpretación de los hechos. Podemos establecerla como una simple *hipótesis de trabajo*, que se plantea como una guía de la investigación o bien con el ánimo de que sea sometida a verificación experimental. En este caso se trata de una *hipótesis científica*. Las hipótesis y los experimentos constituyen los dos polos de la actividad científica (Wartofsky, 1976) y, si queremos clarificar los aspectos metodológicos de la investigación, es preciso precisar, en primer lugar el *significado* de las mismas: “*No hay término científico que padezca mayor ambigüedad que la que aflige a hipótesis: podríamos preparar una lista de enunciados contradictorios acerca de las hipótesis, así como de la condición de que gocen y el uso que se les dé en los estudios científicos*”, (Wartofsky, 1976, p. 241). Así, según este autor, las hipótesis pueden ser: generalizaciones provenientes de la experiencia; inferencias deductivas que realizamos a partir de premisas de orden superior; postulados o suposiciones creados libremente por la inteligencia para poder ordenar o derivar de ellos otros enunciados; intuiciones referentes a lo que parezca evidente o ineludible de creer, o bien a lo que se presente como plausible de alguna manera vaga pero insistente. También pueden entenderse distintos tipos de hipótesis según el alcance y el género de las mismas.

El concepto de *hipótesis*, aparentemente fácil, no se ha revelado como tal en nuestra investigación en donde hemos encontrado muchas dificultades y errores de aprendizaje, así como unas concepciones muy restrictivas sobre el tipo de hipótesis admisibles, especialmente en las aplicaciones (Vallecillos, 1997), lo que afecta de manera muy concreta a la investigación educativa.

En el contexto de la investigación educativa las hipótesis que se plantean son muy frecuentemente *hipótesis estadísticas*, esto es, afirmaciones que se refieren a parámetros de poblaciones. Las hipótesis estadísticas no tienen que ser necesariamente hipótesis científicas en su más amplia acepción, como hemos mencionado antes, si bien pueden considerarse una consecuencia lógica de ellas. Las hipótesis científicas son afirmaciones teóricas sobre poblaciones generalmente no existentes, porque sus elementos no están acotados en el espacio o en el tiempo y, por tanto, no existen en un lugar o momento dado. Sin embargo, las hipótesis estadísticas se refieren siempre a poblaciones existentes, esto es, a poblaciones de las que disponemos de muestras, (Henkel, 1976).

Las hipótesis de investigación son frecuentemente malinterpretadas, insuficientemente determinadas o vagamente expresadas. En palabras de Fox (1981), “*El tema de las relaciones entre las hipótesis y la estadística suele producir gran confusión entre los investigadores principiantes y esa confusión se suele producir, fundamentalmente, porque no se distingue con suficiente claridad entre las hipótesis de investigación y las hipótesis estadísticas Una hipótesis de investigación es la predicción que hace el investigador de una conclusión concreta ... Puede hacer una hipótesis de no diferencia ... En otros casos el investigador espera encontrar alguna diferencia, y esto se reflejará en su hipótesis de investigación ... Pero, ... sea cual fuere el método inferencial que elijamos, sólo podemos contrastar una hipótesis estadística de no diferencia ... [Esta] hipótesis estadística no se expresa formalmente del mismo modo como aparece realmente la hipótesis de investigación en la propuesta o el informe de un estudio. Basta con que el investigador se dé cuenta y comprenda que la hipótesis que está contrastando con el método inferencial es ésta. Si no lo entiende así se pueden producir problemas graves de interpretación*” (pág. 293).

B) Obtención de datos

El proceso de obtención de datos es clave en toda investigación experimental. Usualmente en la investigación educativa las muestras empleadas no son aleatorias, por imposibilidad material de que así sea, de manera que raramente podemos hablar de diseños experimentales que permitan generalizaciones controladas a las poblaciones correspondientes, sino cuasi-experimentales. Si no podemos disponer de muestras aleatorias es necesario por lo menos emplear muestras ‘representativas’ de las poblaciones en

estudio y respetar muy mucho las limitaciones de los resultados obtenidos que no serán generalizables sino a lo sumo descriptivos y orientativos, dependiendo de otras características del diseño como, por ejemplo, el tamaño de las muestras.

Población y muestra

Es conveniente la realización de una prueba piloto con el fin de poner a prueba aspectos concretos de la investigación como, por ejemplo el cuestionario. Se empleará para ello una muestra piloto y se analizarán los resultados obtenidos. Este ensayo nos permite comprobar aspectos técnicos de la prueba como el tiempo necesario para su desarrollo, legibilidad del cuestionario, posibles dificultades para su aplicación en el aula, etc. El análisis de los resultados nos permitirá hacer una clasificación inicial de las variables, la categorización para las variables cualitativas así como el estudio de la fiabilidad y validez de la prueba. Para esto último ver, por ejemplo, Thorndike (1989). Con todos estos datos se determinan los criterios de revisión de la prueba piloto, se redacta la prueba final y se deciden todos los detalles para su aplicación en el aula.

Una vez terminada esta fase se pasa a la fase experimental con la realización de la prueba final con una muestra experimental, a ser posible, aleatoria.

Instrumentos de recogida de datos

Enumeramos los instrumentos empleados en Vallecillos (1994):

1. Prueba escrita general: cuestionario sobre aspectos conceptuales y un problema de aplicación de un test de hipótesis.
 - 1.1. Cuestionario sobre aspectos conceptuales: consta de 20 ítems, 5 de ellos de Verdadero / Falso y 15 con 4 opciones y sólo una correcta.
 - 1.1.1. Opción elegida.
 - 1.1.2. Argumento de apoyo: en 6 de los 20 ítems se pedía razonar la respuesta.
 - 1.2. Prueba de ensayo sobre contenidos procedimentales: se propuso un problema de aplicación con 4 preguntas.
2. Entrevista personal: a siete estudiantes seleccionados intencionalmente.
 - 2.1. Cuestionario escrito: 3 ítems y 2 problemas con tres preguntas cada uno.
 - 2.2. Explicación oral del trabajo realizado.

C) Análisis de los resultados

En nuestro caso, se realizó un primer análisis descriptivo de tipo cuantitativo, por titulaciones y global, de los resultados de tipo conceptual para cada uno de los 20 ítems de la prueba. Se determinaron así los principales errores y dificultades de aprendizaje de los estudiantes de la muestra sobre el aprendizaje del contraste de hipótesis observados en forma sistémica y no aisladamente.

Se realizó también un análisis de tipo cualitativo de las argumentaciones proporcionadas por los alumnos en los seis ítems en los que se les pidió que razonaran la respuesta y un análisis conjunto de respuestas y argumentos.

Finalmente, se realizó el análisis de la parte procedimental de la prueba y de la influencia de los errores de tipo conceptual en la resolución del problema propuesto.

Las conclusiones obtenidas se organizan en cuatro grandes bloques: errores conceptuales asociados a elementos concretos de significado, estructura de las respuestas al cuestionario, concepciones sobre conceptos básicos en el contraste de hipótesis y conocimiento procedimental.

Posteriormente se llevó a cabo una entrevista con un grupo seleccionado de siete estudiantes cuya metodología, objetivos e hipótesis fueron también previamente determinados. Los principales resultados de ella pueden verse en Vallecillos y Batanero (1997).

UN EJEMPLO: COMPRENSIÓN DEL NIVEL DE SIGNIFICACIÓN

A continuación describimos los datos relativos al estudio de la comprensión del concepto de nivel de significación en un test de hipótesis obtenidos en Vallecillos (1994), que es un concepto clave en este procedimiento. Pueden verse también Vallecillos (1998a; 1998b; en revisión) y Vallecillos y Batanero (1996a; 1996b; 1997).

Objetivos

En síntesis, los principales objetivos de la investigación referidos al concepto de nivel de significación son los siguientes:

- Analizar en forma sistémica el aprendizaje de los tests de hipótesis por estudiantes universitarios.
- Obtener evidencias empíricas acerca del aprendizaje del concepto de nivel de significación.
- Identificar dificultades de aprendizaje, interpretaciones incorrectas y las concepciones de los estudiantes acerca del concepto de nivel de significación.

Obtención de datos: cuestionario

*Ítem 1: Un nivel de significación del 5% significa que, en promedio, 5 de cada 100 veces que rechazamos la hipótesis nula estaremos equivocados: $V / *F$*

Razona la respuesta.

*Ítem 2: Un nivel de significación del 5% significa que, en promedio, 5 de cada 100 veces que la hipótesis nula es cierta la rechazaremos: $*V / F$*

Razona la respuesta.

[Nota: con * la respuesta correcta]

Los dos ítems son sintácticamente muy parecidos, sin embargo, son completamente diferentes, el segundo de ellos es la definición del concepto de nivel de significación mientras que el primero no lo es. Con la comparación de estas dos expresiones verbales, pretendemos analizar el error de confusión de los dos sucesos, condicional y condicionado, que intervienen en la definición, descrito anteriormente por Birnbaum (1982) y Falk (1986).

Resultados

Comentamos en primer lugar los resultados obtenidos como respuesta a cada uno de los ítems separadamente y después los argumentos que los alumnos han dado como explicación de su respuesta. A continuación llevamos a cabo un análisis conjunto de los dos ítems con el propósito de detectar la coherencia entre las dos respuestas y profundizar así en la comprensión del concepto de nivel de significación los alumnos participantes en la investigación.

Respuestas a los ítems

Los resultados están contenidos en la Tabla 1 a continuación.

En el ítem 1 la respuesta mayoritaria, de más de la mitad de los estudiantes, el 53%, es incorrecta. En el ítem 2 la respuesta mayoritaria, también de más de la mitad de los estudiantes, 53.7%, es correcta. Sin embargo, los porcentajes de estudiantes que responden correctamente al ítem 2 y al ítem 1 deberían haber sido semejantes. Aparentemente, la respuesta mayoritaria incorrecta en el ítem 1 expresa el error de confusión entre los dos sucesos condicionales que definen el nivel de significación al que nos hemos referido antes, descrito por Birnbaum (1982) y Falk (1986).

Tabla 1: Frecuencia y porcentajes de la opción elegida a los ítems

Ítems	Correcta	Incorrecta	Blanco	Total
Ítem 1	141 (32.3%)	231 (53%)	64 (14.7%)	436
Ítem 2	234 (53.7%)	131 (30%)	71 (16.3%)	436

Creemos que, en ambos casos, el porcentaje de respuestas correctas no es muy alto mientras que si lo es el de respuestas en blanco lo que parece indicar un alto grado de dificultad en la expresión verbal de este concepto. Aunque estos datos son en si mismos muy reveladores, es necesario hacer un análisis más fino para poder entender mejor la comprensión del nivel de significación conseguida por estos alumnos así como sus concepciones al respecto. Analizamos para ello los argumentos que dan como apoyo a sus respuestas a cada uno de los ítems. A continuación incluimos la categorización de argumentos y los resultados obtenidos en las Tablas 2 y 3.

Categorías de argumentos

La categorización de los argumentos empleados por los estudiantes para apoyar su respuesta a los ítems es única ya que ambos se refieren al mismo concepto. Analizamos primero las respuestas dadas en los dos ítems separadamente y después lo hacemos conjuntamente. El proceso de categorización de las respuestas ha sido muy laborioso y pormenorizado. En primer lugar se recogieron con mucho detalle todas las posibles respuestas de los estudiantes que dieron lugar a una primera clasificación de argumentos con muchas categorías. A continuación, por un proceso de refinamiento sucesivo, se fue reduciendo el número de estas clases atendiendo a los aspectos esenciales determinados en el minucioso análisis de contenido del estudio realizado previamente (Vallecillos, 1994). Finalmente quedaron las siete categorías de argumentos que se describen a continuación, de las que se ofrece, además un ejemplo textual de respuesta de un alumno que hemos incluido en ella para ilustración del lector.

1. *Argumento correcto*: por definición del nivel de significación, como probabilidad de rechazar la hipótesis nula siendo cierta, expresada verbalmente o en forma simbólica $\alpha = P(\text{rechazar } H_0 \mid H_0 \text{ cierta})$, empleada en forma consistente con la respuesta al ítem. Por ejemplo, un estudiante responde correctamente al ítem 2 y razona su respuesta con el siguiente argumento:

- "Por la definición del nivel de significación $\alpha = P(\text{rechazar } H_0 \mid H_0 \text{ cierta})$ si $\alpha = 0.05$ significa que rechazaremos la hipótesis nula siendo cierta 5 de cada cien veces".

2. *Argumento correcto*: por definición del nivel de significación, $\alpha = P(\text{rechazar } H_0 \mid H_0 \text{ cierta})$, empleada en forma contradictoria con la respuesta al ítem, esto es, la explicación proporcionada debería haber llevado al estudiante a responder, por ejemplo, Falso, en el ítem y sin embargo ha señalado Verdadero. Por ejemplo, un estudiante da una respuesta incorrecta al ítem 2 y argumenta:

- "Significa rechazar, en promedio, 5 de cada 100 veces la hipótesis nula siendo cierta".

3. *Argumento erróneo*: confusión del nivel de significación α con la probabilidad 'a posteriori' de que la hipótesis nula sea cierta, esto es, con $P(H_0 \text{ cierta} \mid H_0 \text{ rechazada}) \neq \alpha$. Por ejemplo, un estudiante responde incorrectamente el ítem 2 y argumenta:

- “No se trata de rechazar H_0 en un 5% sino de equivocarnos en un 5% cuando la rechazamos”.
- 4. *Argumento erróneo*: confusión en la definición del nivel de significación α pero conservando en ella la idea de probabilidad condicional. Por ejemplo, algunos estudiantes confunden α con la probabilidad de ‘acertar’ una vez que se ha aceptado la hipótesis nula, esto es, con $P(H_0 \text{ cierta} \mid H_0 \text{ aceptada}) \neq \alpha$. Una respuesta de este grupo, que responde correctamente al ítem 1, es la siguiente:
 - “Quiere decir que la probabilidad de que hayamos acertado al aceptar H_0 es del 15%”.
- 5. *Argumento erróneo*: confusión en la definición de α considerándola como una probabilidad no condicional referida a la hipótesis nula, independientemente de que se acepte o se rechace o se la considere cierta o no. Por ejemplo, unos estudiantes consideran α como probabilidad simple de aceptar la hipótesis nula y otros de rechazarla. Una respuesta de esta categoría, de un estudiante que responde incorrectamente al ítem 2, es la siguiente:
 - “Significa que hay una probabilidad del 5% de que H_0 sea rechazada”.
- 6. *Argumento erróneo*: confusión de α con la probabilidad de error o bien como probabilidad de error en la decisión tomada, sea cual sea ésta y sea cierta o no la hipótesis nula. Por ejemplo, un argumento de un estudiante que responde incorrectamente al ítem 2, es el siguiente:
 - “Lo que significa es que 5 de cada 100 veces nos equivocaremos en la hipótesis aceptada”.
- 7. *Argumentos difícilmente interpretables*. Están clasificadas aquí las argumentaciones confusas o inoportunas, esto es, que no responden directamente a la cuestión propuesta aunque se refieran a conceptos relativos al contraste de hipótesis.

Análisis de argumentos

En las Tabla 2 están contenidos los resultados obtenidos en el ítem 1 y en la Tabla 3 los correspondientes al ítem 2.

Una primera cuestión llamativa es el hecho de que, de los 436 estudiantes que han respondido el ítem 1, sólo 191 (43.8%) han proporcionado una argumentación explicativa de su respuesta. Esto nos da un 56.2% de respuestas en blanco, que indica un alto grado de dificultad para los estudiantes en la realización de esta tarea.

Un 47.1% en total de los argumentos proporcionados por los estudiantes son correctos, si bien muchos de ellos han argumentado en forma contradictoria con su respuesta al ítem.

Tabla 2: Frecuencias y porcentajes de argumentos en el ítem 1

Frecuencia % Fila	ITEM 1: ARGUMENTOS							Total Fila
	1 (C)	2 (C)	3 (E)	4(E)	5 (E)	6 (E)	7 (DI)	
Camino	6 15.4	11 28.2	2 5.1	8 20.5	2 5.1	6 15.4	4 10.3	39 20.4
Informática	2 25.0	2 25.0		2 25.0	1 12.5		1 12.5	8 4.2
Pedagogía		3 14.3	1 4.8	4 19.0	4 19.0	7 33.3	2 9.5	21 11.0
Psicología	5 14.3	11 31.4	2 5.7	8 22.9		6 17.1	3 8.6	35 18.3
Medicina	10 27.8	10 27.8	6 16.7	3 8.3	2 5.6	2 5.6	3 8.3	36 18.8
Matemáticas	7 31.8	3 13.6		3 13.6	5 22.7	1 4.5	3 13.6	22 11.5
Empresariales	4 13.3	16 53.3	1 3.3	5 16.7	1 3.3		3 10.0	30 15.7
Total Col.	34 17.8	56 29.3	12 6.3	33 17.3	15 7.9	22 11.5	19 9.9	191 100

C: Correcto

E: Erróneo

DI: Difícil de interpretar

En la columna 3, en donde están clasificadas las respuestas que explicitan la confusión de α con la probabilidad 'a posteriori' de que la hipótesis nula sea cierta una vez que ha sido rechazada, esto es, el intercambio entre los dos sucesos que intervienen en la probabilidad condicional que define el nivel de significación, $P(H_0 \text{ cierta} \mid H_0 \text{ rechazada})$, aparecen porcentajes bajos en general en todos los grupos de estudiantes, y un 6.3% para el total. Nos fijamos en este tipo de error porque es del que tenemos referencias anteriores en la literatura de investigación, Birnbaum (1982), Falk (1986). Sin embargo, en nuestro estudio ha aparecido relativamente poco, como hemos visto.

El error más extendido entre nuestros estudiantes ha sido el de confundir α con la probabilidad de 'acertar' una vez que se ha aceptado la hipótesis nula, esto es, con $P(H_0 \text{ cierta} \mid H_0 \text{ aceptada}) \neq \alpha$. Los resultados incluidos en la columna 4 de la Tabla 2, suponen un 17.3% de las respuestas de los estudiantes. Estos, aún confundiendo la definición del nivel de significación, conservan sin embargo una característica importante de la misma que es la idea de probabilidad condicional.

El resto de los argumentos erróneos proporcionados por los estudiantes se dan en modestos porcentajes lo que nos indica que tienen una incidencia baja.

En la Tabla 3, están contenidos los resultados correspondientes al ítem 2.

También aquí han abundado las respuestas en blanco, 65.4%, ya que sólo han argumentado su respuesta 151 estudiantes, un 34.6% de los estudiantes que han respondido el ítem 2.

Tabla 3: Frecuencias y porcentajes de argumentos en el ítem 2

Frecuencia % fila	ITEM 2: ARGUMENTOS							Total fila
	1 (C)	2 (C)	3 (E)	4 (E)	5 (E)	6 (E)	7 (DI)	
Camino.	14 38.9	5 13.9	2 5.6	5 13.9	3 8.3	4 11.1	3 8.3	36 23.8
Informática		2 50.0		1 25.0			1 25.0	4 2.6
Pedagogía		1 6.3		1 6.3	3 8.8	8 50.0	3 18.8	16 10.6
Psicología	10 45.5	2 9.1	2 9.1	5 22.7		2 9.1	1 4.5	22 14.6
Medicina	15 42.9	8 22.9	3 8.6	2 5.7	1 2.9	3 8.6	3 8.6	35 23.2
Matemáticas	8 44.4	2 11.1	1 5.6	2 11.1	4 22.2	1 5.6		18 11.9
Empresariales	9 45.0	3 15.0	1 5.0	2 10.0	5 25.0			20 13.2
Total columna	56 37.1	23 15.2	9 6.0	18 11.9	16 10.6	18 11.9	11 7.3	151 100

C: Correcto

E: Erróneo

DI: Difícil de interpretar

Un 52.3% de los estudiantes ha respondido correctamente al ítem 2, porcentaje ligeramente superior al obtenido en el ítem 1. También aquí el mayor porcentaje obtenido corresponde al mismo argumento erróneo que en el ítem anterior, columna 4 de la Tabla 3, el que expresa la confusión de α con la probabilidad de 'acertar' una vez que se ha aceptado la hipótesis nula, si bien en menor medida, un 11.9% frente 17.3 % anterior.

En lo que se refiere al error de confusión de los condicionales en la definición de α , argumento codificado como 3, también el porcentaje es bajo, 6%, lo que confirma el resultado obtenido en el ítem 1.

En resumen, con mayor o menor incidencia entre los estudiantes de la muestra, sus argumentaciones expresan principalmente los siguientes errores:

Principales errores encontrados:

- Muchos estudiantes confunden a con la probabilidad de ‘acertar’ en caso de aceptar la hipótesis nula, esto es, de que ésta sea cierta una vez que ha sido aceptada. Este es el error más frecuente encontrado en las respuestas de los estudiantes de nuestra muestra.
- Muchos estudiantes confunden el concepto de nivel de significación a con la probabilidad ‘a posteriori’ de la hipótesis nula.
- Muchos estudiantes responden simultáneamente verdadero o falso a ambas cuestiones, esto es, son incoherentes. La expresión verbal de la definición de a puede resultar demasiado difícil de interpretar para ellos.
- Muchos estudiantes confunden a con una probabilidad no condicional referida a la hipótesis nula.
- Muchos estudiantes confunden a con una probabilidad de error, cualquiera que sea la decisión tomada y como quiera que sea la hipótesis, verdadera o falsa.

Análisis conjunto de argumentos

Aunque el número de respuestas no es muy grande (191 en el ítem 1, 151 en el ítem 2 y 135 en ambos), una vez analizados separadamente los resultados obtenidos en los dos ítems propuestos, hemos comparado ambos resultados con el fin de estudiar la consistencia de las respuestas a ambos ítems realizando una tabulación cruzada que se recoge en la Tabla 4.

Explicitan el mismo tipo de error de intercambio de los dos sucesos que intervienen en la definición del nivel de significación, argumento codificado como 3, en ambos ítems sólo 3 estudiantes. Estos son los, que al mantener de forma explícita, el mismo tipo de argumentación errónea en ambos ítems podemos considerar ‘estables’ en su manifestación. Separadamente, por ítems suponen 3 de los 12 que había en el ítem 1 y de los 9 que hubo en el ítem 2, 25% y 33.3% respectivamente. Estos resultados, pues, no señalan este error como el más significativo entre los que afectan al aprendizaje del concepto de nivel de significación.

Se manifiesta con una mayor estabilidad la argumentación errónea codificada como 4, que expresa la confusión de a con la probabilidad de ‘acertar’ una vez que se ha aceptado la hipótesis nula, esto es, la probabilidad ‘a posteriori’ de la verdad de la hipótesis nula aceptada, aún manteniendo la idea de probabilidad condicional para ella. Se han obtenido 10 respuestas, que suponen un 30.3% y un 55.5% de las obtenidas, respectivamente, en los ítems 1 y 2.

En total, sumando todas las frecuencias de las casillas correspondientes en la diagonal principal de la Tabla 4, obtenemos 55 estudiantes que dan argumentaciones correctas en ambos ítems y 31 estudiantes que manifiestan un mismo tipo de error en ellas, esto es, que manifiestan un tipo de argumentación errónea estable. Hay 29 estudiantes que dan una argumentación correcta en uno de los ítems e incorrecta en el otro y 20 que manifiestan distintas argumentaciones erróneas en uno y otro ítem, esto es, al menos aparentemente, son inestables en sus argumentaciones. Esta situación nos ha llevado a realizar otra tabulación cruzada de estas argumentaciones.

Tabla 4: Tabla cruzada de argumentos en los ítems 1 y 2

Frecuencia % de Fila % de Columna		ARGUMENTOS EN EL ÍTEM 2						Total Fila	
		1 (C)	2 (C)	3 (E)	4 (E)	5 (E)	6 (E)		7 (DI)
ARGUMENTOS EN EL ÍTEM 1	1 (C)	19 76.0 38.0	1 4.0 5.0	2 8.0 28.6	2 8.0 11.8			1 4.0 11.1	25 18.5
	2 (C)	20 45.5 40.0	15 34.1 75.0	1 2.3 14.3	2 4.5 11.8	3 6.8 21.4	2 4.5 11.8	1 2.3 11.8	44 32.6
	3 (E)		2 18.2 10.0	3 27.3 42.9	2 18.2 11.8	2 18.2 14.3	2 18.2 11.1		11 8.1
	4 (E)	4 19.0 8.0	1 4.8 5.0		10 47.6 58.8	1 4.8 7.1	4 19.0 22.2	1 4.8 11.1	21 15.6
	5 (E)	3 25.0 6.0			1 8.3 5.9	5 41.7 35.7	2 16.7 11.1	1 8.3 11.1	12 8.9
	6 (E)	2 15.4 4.0		1 7.7 14.3		2 15.4 14.3	8 61.5 44.4		13 9.6
	7 (DI)	2 22.2 4.0	1 11.1 5.0			1 11.1 7.1		5 55.6 55.6	9 6.7
Total Columna		50 37.0	20 14.8	7 5.2	17 12.6	14 10.4	18 13.3	9 6.7	135 100

C: Correcto

E: Erróneo

DI: Difícil de interpretar

Podemos observar mejor la estabilidad de las respuestas de los estudiantes cuando estas se clasifican simplemente en correctas o no. En la Tabla 5 podemos ver los resultados obtenidos:

55 estudiantes, un 40.7% del total de los que responden, dan una argumentación correcta en ambos ítems; 51 estudiantes, un 37.7% del total, lo hacen incorrectamente en los dos ítems, aunque manifestando diversos tipos de error y 29, un 21.5% del total responden correctamente en uno de ellos e incorrectamente en el otro.

Tabla 5: Tabla reducida de argumentos en los ítems 1 y 2

Frecuencia % de Fila % de Columna		ÍTEM 2: ARGUMENTOS		Total fila
		Correcto	Erróneo	
ÍTEM 1: ARGUMENTOS	Correcto	55 79.7 78.6	14 20.3 21.5	69 51.1
	Erróneo	15 22.7 21.4	51 77.3 78.5	66 48.9
Total Columna		70 51.9	65 48.1	135 100.0

Interpretaciones del nivel de significación

Los argumentos dados por los estudiantes nos han permitido conocer las distintas interpretaciones que éstos hacen del concepto de nivel de significación. El análisis de estos argumentos que venimos comentando nos permite clasificar, a modo de resumen, los resultados obtenidos en cuanto a estas distintas interpretaciones del concepto de nivel de significación que manifiestan los estudiantes en sus respuestas. Las hemos separado en tres grandes grupos: correctas, incorrectas pero conservando la idea de probabilidad condicional e incorrectas, asociadas a la probabilidad de error en distintas formas. En la Tabla 6 están contenidas las frecuencias obtenidas en las distintas interpretaciones en los dos ítems analizados.

a) Interpretaciones de α correctas

Son aquellas en las que se interpreta el nivel de significación como probabilidad de rechazar la hipótesis nula siendo cierta. Hemos encontrado dos variantes:

I_1 : Se emplea la definición en forma consistente con la respuesta (V/F) al ítem.

I_2 : Se emplea la definición en forma contradictoria con la respuesta (V/F) al ítem.

b) Interpretaciones de α incorrectas pero conservando la idea de probabilidad condicional

Hemos encontrado las siguientes:

I_3 : Nivel de significación como probabilidad de que la hipótesis nula sea cierta una vez que se ha rechazado, esto es, como probabilidad de error en la decisión en caso de rechazo.

I_4 : Confusión del nivel de significación con la probabilidad de haber 'acertado' al aceptar la hipótesis nula.

I_5 : Nivel de significación como probabilidad de que la hipótesis nula sea falsa una vez que se ha aceptado, esto es, como probabilidad de error en la decisión en caso de aceptación.

I_6 : Nivel de significación como probabilidad de que la hipótesis nula sea falsa una vez que se ha rechazado, esto es, como probabilidad de acierto en la decisión en caso de rechazo.

I_7 : Nivel de significación como probabilidad de aceptar la hipótesis nula siendo falsa, esto es, confusión con el error de tipo II, β .

c) Interpretaciones de α como probabilidad no condicional referida a la hipótesis nula

I_8 : Nivel de significación como probabilidad simple de aceptar la hipótesis nula.

I_9 : Nivel de significación asociado con el rechazo de la hipótesis nula falsa.

I_{10} : Nivel de significación como probabilidad de que la hipótesis nula sea falsa.

I_{11} : Nivel de significación como probabilidad de que la hipótesis nula sea cierta.

Tabla 6: Frecuencias de interpretaciones del nivel de significación

INTERPRETACIONES	ITEM 1	ITEM 2
I ₁ : $\alpha = P(\text{rechazar } H_0 \mid H_0 \text{ cierta})$ (1)	34	56
I ₂ : $\alpha = P(\text{rechazar } H_0 \mid H_0 \text{ cierta})$ (2)	56	23
I ₃ : $\alpha = P(H_0 \text{ cierta} \mid H_0 \text{ rechazada})$	12	9
I ₄ : $\alpha = P(H_0 \text{ cierta} \mid H_0 \text{ aceptada})$	3	3
I ₅ : $\alpha = P(H_0 \text{ falsa} \mid H_0 \text{ aceptada})$	22	3
I ₆ : $\alpha = P(H_0 \text{ falsa} \mid H_0 \text{ rechazada})$	2	1
I ₇ : $\alpha = P(\text{aceptar } H_0 \mid H_0 \text{ falsa})$	6	11
I ₈ : $\alpha = P(\text{aceptar } H_0)$	1	1
I ₉ : $\alpha = P(\text{rechazar } H_0 \text{ falsa})$	10	8
I ₁₀ : $\alpha = P(H_0 \text{ falsa})$	1	1
I ₁₁ : $\alpha = P(H_0 \text{ cierta})$	3	6
I ₁₂ : $\alpha = P(\text{error en la decisión})$	11	12
I ₁₃ : $\alpha = P(\text{error})$	10	6
I ₁₄ : Confusión de α con $1 - \alpha$	1	
I ₁₅ : Otras interpretaciones	19	11
Total	191	151

(1): Consistente con la opción elegida

(2): Inconsistente con la opción elegida

d) Interpretaciones de α como probabilidad de error

I₁₂: Nivel de significación como probabilidad simple de error en la decisión tomada.

I₁₃: Nivel de significación como probabilidad de error.

e) Otras interpretaciones

I₁₄: Confusión de α y $1 - \alpha$

I₁₅: Interpretaciones confusas o inoportunas.

Concepciones sobre el concepto de nivel de significación

Consideramos que estas distintas interpretaciones que los estudiantes hacen del concepto de nivel de significación, son la manifestación de las concepciones (Artigue, 1990) que estos tienen sobre el correspondiente concepto, que podemos sintetizar como sigue:

Concepción correcta del nivel de significación, como probabilidad condicional de rechazar la hipótesis nula siendo cierta, si bien en algunos casos se aplica en forma contradictoria en la respuesta (V/F) del ítem

correspondiente. Es la concepción más frecuente con un total de 90 estudiantes, 47.1% del total de respuestas en el ítem 1, y 79, 52.3% del total de respuestas en el ítem 2. Incluimos aquí las interpretaciones I_1 e I_2 .

Además de esta concepción correcta del nivel de significación, hemos encontrado tres más incorrectas en las argumentaciones de los estudiantes que han tomado parte en nuestra investigación:

Nivel de significación como probabilidad condicional referida a una de las hipótesis: Incluimos aquí a los estudiantes que conservan la idea de probabilidad condicional pero intercambian los sucesos condicionado o condicionante o ambos. Aquí están incluidos los que manifiestan la confusión previamente descrita por (Falk, 1986) como caso particular. Incluimos las interpretaciones de I_3 a I_7 . Tenemos 45 casos, 23.6% del total de respuestas en el ítem 1 y 27, el 17.9% del total de respuestas en el ítem 2, 72 en total.

Nivel de significación como probabilidad simple de la hipótesis nula: Muchos estudiantes interpretan el nivel de significación como una probabilidad no condicional de la hipótesis nula, tanto en el caso de aceptación como de rechazo. Incluimos las interpretaciones de I_8 a I_{11} . Hemos encontrado 15 casos, 7.9% del total de repuestas en el ítem 1 y 16 casos, 10.6% de las respuestas en el ítem 2, 31 en total.

Nivel de significación como probabilidad de error: Incluimos aquí las respuestas que se refieren a a como la probabilidad de error, tanto si se asocia a la hipótesis nula como a la alternativa y tanto si ésta es aceptada como rechazada. Incluimos las interpretaciones I_{12} e I_{13} . Hemos obtenido 21 casos, 11% en el ítem 1 y 18, 11.9% en el ítem 2.

En resumen, hemos encontrado una gran variedad de interpretaciones del nivel de significación en las respuestas de los estudiantes, la mayor parte de ellas incorrectas. Estas manifestaciones nos permiten aproximarnos a las concepciones que tienen del concepto de nivel de significación. A continuación resumimos las principales concepciones incorrectas encontradas en el análisis de los argumentos descritos anteriormente:

Principales concepciones erróneas sobre el concepto de nivel de significación:

- El nivel de significación, α como probabilidad ‘a posteriori’ de la hipótesis nula o la alternativa.
- El nivel de significación, α como probabilidad de que la hipótesis nula sea verdadera una vez que ha sido rechazada.
- El nivel de significación, α como probabilidad condicional ya sea de la hipótesis nula o de la alternativa.
- El nivel de significación, α como probabilidad no condicional de la hipótesis nula.
- El nivel de significación, α como la probabilidad de error en un test de hipótesis.

EL INVESTIGADOR Y EL TÉCNICO ESTADÍSTICO

Por último, creemos conveniente llamar la atención de los investigadores que trabajan en el campo de la educación sobre un aspecto no siempre tenido en cuenta. Con frecuencia, la investigación educativa comporta el análisis de datos y el uso de técnicas estadísticas fuera del interés propio del investigador y que requieren el concurso técnico de un profesional estadístico. En este caso el colaborador estadístico deberá ‘comprender’ bien el problema de investigación, buscar la solución técnica más adecuada y proporcionar la respuesta en un lenguaje ‘inteligible’ al investigador. Sin embargo, aún en las mejores condiciones de comunicación entre el investigador y el técnico estadístico, el problema de investigación sólo tiene sentido completo para el investigador que es el que tiene que dar respuesta, con la ayuda técnica precisa, a ‘su’ problema de investigación. Una reflexión sobre el tema puede verse en Batanero, Godino y Vallecillos (1992).

CONCLUSIONES

Sin pretender elevar a categoría la metodología empleada en la investigación de referencia y otras realizadas posteriormente con un esquema metodológico basado total o parcialmente en el descrito, hemos comprobado que ésta puede ser muy útil para la obtención de resultados en el campo concreto de la investigación en Educación Matemática. Ha resultado adecuada para la consecución de los objetivos previstos y los múltiples procedimientos de obtención de los datos, las variables analizadas, los análisis estadísticos empleados y los resultados obtenidos permiten considerarla útil para la investigación de procesos de enseñanza y aprendizaje en otros contextos y problemas educativos.

Queremos terminar haciendo una llamada a la sensatez y la honestidad en la investigación educativa utilizando palabras de Fox (1981): "... para que la investigación en ciencias sociales alcance resultados significativos, tiene que ser un proceso racional. Tenemos que saber por qué hacemos lo que hacemos, y por qué lo hacemos como lo hacemos, y también tenemos que saber qué formas de hacerlo hemos tenido en cuenta y cuáles hemos rechazado y por qué", (pág. 23).

AGRADECIMIENTOS

A los Proyectos de Investigación PB97-0827 y BS02000-1507, financiados por el Ministerio de Ciencia y Tecnología, Madrid, España.

REFERENCIAS

- Afifi, A. A. y Azen, S. P. (1990). *Computer-aided multivariable analysis*. New York: Van Nostrand Co.
- Artigue, M. (1990). Épistémologie et Didactique. *Recherches en Didactique des Mathématiques*, 10(2-3), 241-286.
- Batanero, C.; Godino, J. D. y Vallecillos, A. (1992). El análisis de datos como útil y como objeto en Didáctica de la Matemática. *Educación Matemática*, 4(1), 46-53.
- Birnbaum, I. (1982). Interpreting Statistical Significance. *Teaching Statistics*, 4, 24-27.
- Bisquerra, R. (1989). *Introducción conceptual al análisis multivariable. Un enfoque informático con los paquetes SPSS-X, BMDP, LISREL y SPAD*. Barcelona: PPU.
- Falk, R. (1986). Misconceptions of Statistical Significance. *Journal of Structural Learning*, 9, 83-96.
- Fox, J. D. (1981). *El proceso de investigación en educación*. Pamplona: Editorial Universidad de Navarra, S. A.
- Godino, J. D. y Batanero, C. (1994). Significado institucional y personal de los objetos matemáticos. *Recherches en Didactique des Mathématiques*, 14(3), 325-355.
- Henkel, R. (1976). *Test of Significance*. Beverly Hills: Sage Pub.
- Thorndike, E. L. (1989). *Psicometría Aplicada*. México: Limusa.
- Vallecillos, A. (1994). *Estudio teórico-experimental de errores y concepciones sobre el contraste estadístico de hipótesis en estudiantes universitarios*. Tesis Doctoral. Universidad de Granada.
- Vallecillos, A. (1995). Comprensión de la lógica del contraste de hipótesis en estudiantes universitarios. *Recherches en Didactique des Mathématiques*, 15(3), 53-81.
- Vallecillos, A. (1996a). Inferencia estadística y enseñanza: un análisis didáctico del contraste de hipótesis estadísticas. Granada: Comares.
- Vallecillos, A. (1996b). Students' conceptions of the logic of hypothesis testing. *Hiroshima Journal of Mathematics Education*, Vol. 4, 43-61.

- Vallecillos, A. (1977). El papel de las hipótesis estadísticas en los contrastes: concepciones y dificultades de aprendizaje. *Educación Matemática*, 9(2), 5-20.
- Vallecillos, A. (1998a) Experimental study on the learning of the significance level concept. In: Pereira-Mendoza, L.; Lua Seu Kea; Tang Wee Kee & Wing-Keung Wong (Eds.): *Proceeding of the Fifth International Conference on Teaching of Statistics*, Vol. 3, 1475-1476. Singapore: ISI.
- Vallecillos, A. (1998b). Research and Teaching of Statistical Inference. In: International Conference on the Teaching of Mathematics, pp. 296-298. John Wiley & Sons, Inc. Publishers.
- Vallecillos, A. (1999). Some empirical evidences on learning difficulties about testing hypotheses. Ponencia Invitada, Topic IPM 58. Proceedings of the 52nd Session of the International Statistical Institute, Vol. 2, Tome LVIII, pp. 201-204. The Netherlands: ISI.
- Vallecillos, A. (en revisión). Learning from experience: the level of significance concept. *Hiroshima Journal of Mathematics Education*.
- Vallecillos, A. y Batanero, C. (1996a). Conditional probability and the level of significance in the tests of hypotheses. In: Gutiérrez, A. y Puig, L. (Eds.): *Proceedings of the 20 PME*, Vol. 4, pp. 371-378. Valencia.
- Vallecillos, A. y Batanero, C. (1996b). Factors affecting student' interpretation of the significance level in tests of hypotheses. *Proceedings of the ICME-8*, p. 135. Sevilla.
- Vallecillos, A. y Batanero, C. (1997). Conceptos activados en el contraste de hipótesis estadísticas y su comprensión por estudiantes universitarios. *Recherches en Didactique des Mathématiques*, 17(1), 29-48.
- Wartofsky, M. W. (1976). *Introducción a la filosofía de la Ciencia*. Madrid: Alianza Universidad.