

Dificultades en la construcción de intervalos de confianza por estudiantes de Bachillerato y de Psicología

Carmen Batanero¹

batanero@ugr.es

<https://orcid.org/0000-0002-4189-7139>

Antonio Francisco Roldán López de Hierro¹

aroldan@ugr.es

<http://orcid.org/0000-0002-6956-4328>

Rocío Álvarez Arroyo¹

rocioaarroyo@ugr.es

<https://orcid.org/0000-0002-3201-8542>

¹Universidad de Granada

Granada, España.

Recibido: 21/01/2020 **Aceptado:** 15/03/2020

Resumen

El intervalo de confianza es un procedimiento básico en inferencia estadística y, por ello, su estudio se incluye en las Matemáticas Aplicadas a las Ciencias Sociales II para alumnos de dicha modalidad de Bachillerato y en el primer curso de la licenciatura de Psicología. Además, en las pruebas de acceso a la universidad se propone con frecuencia un problema relacionado con este contenido. Con el objetivo de evaluar la dificultad que implica la construcción del intervalo, en este trabajo se analizan los pasos requeridos para resolver un problema abierto tomado de las pruebas españolas de acceso a la universidad y se comparan las soluciones aportadas por 58 estudiantes de Bachillerato y 57 de Psicología, después de haber estudiado el tema. Los resultados muestran una mayor proporción de respuestas correctas en estos últimos estudiantes, que han seguido un método ligeramente diferente a sus compañeros, que parece ser más comprensible para ellos. Se informa también de la frecuencia con que se completan diferentes pasos y de los principales errores en el proceso.

Palabras clave: Intervalo de confianza, Construcción, Estudiantes de Bachillerato, Estudiantes de Psicología

Dificuldades dos alunos do ensino médio e de psicologia na construção de intervalos de confiança

Resumo

O intervalo de confiança é um procedimento básico na inferência estatística e, portanto, seu estudo está incluído na Matemática Aplicada às Ciências Sociais II para estudantes da referida modalidade de Bacharelado e no primeiro ano do curso de Psicologia. Além disso, nos exames de admissão na universidade é frequentemente proposto um problema relacionado a esse conteúdo. Com o objetivo de avaliar a dificuldade envolvida na construção do intervalo, este trabalho analisa as etapas necessárias para solucionar um problema em aberto realizado nos exames de ingresso na universidade espanhola e compara as soluções fornecidas por 58 alunos do ensino médio e 57 do ensino médio. Psicologia, depois de ter estudado o assunto. Os

resultados mostram uma proporção maior de respostas corretas nesses últimos alunos, que seguiram um método ligeiramente diferente do de seus colegas, o que parece ser mais compreensível para eles. A frequência com que etapas diferentes são concluídas e os principais erros no processo também são relatados.

Palavras-chave: Intervalo de confiança, Construção, Estudantes de Bacharelado, Estudantes de Psicologia

High school and Psychology students' difficulties in building confidence intervals

Abstract

Confidence interval is a basic procedure in statistical inference and, therefore, its study is included in the Mathematics Applied to Social Sciences II for high school students and in the first course of the Psychology degree. In addition, a problem related to this content is frequently proposed at the university entrance exams. In order to evaluate the difficulty involved in the construction of the interval, in this work the steps required to solve an open problem taken from the Spanish university entrance exams are analysed and the solutions provided by 58 high school students and 57 Psychology students are compared after they studied the topic. The results show a higher proportion of correct answers in these last students, who have followed a slightly different method from their peers, which seems to be more understandable to them. The frequency with which different steps are completed and the main errors in the process are also reported.

Keywords: Confidence interval, Construction, High school students, Psychology students.

Introducción

El intervalo de confianza es un procedimiento básico en inferencia estadística que se utiliza como complemento (o en sustitución) de los contrastes de hipótesis. Entidades profesionales, como la American Psychological Association (APA), promueven actualmente su uso, que consideran más intuitivo y más informativo que los contrastes (Coulson, Healey, Fidler y Cumming, 2010; Wilkinson y TFSA, 1999; Yaremko, Harari, Harrison y Lynn, 2013).

En España su enseñanza se incluye en el Bachillerato de Ciencias Sociales (MECD, 2015), con los siguientes contenidos: estimación por intervalos de confianza; relación entre confianza, error y tamaño muestral; intervalo de confianza para la media poblacional de una distribución normal con desviación típica conocida; intervalo de confianza para la media poblacional de una distribución de modelo desconocido y para la proporción en el caso de muestras grandes. Por otro lado, en las pruebas de acceso a la universidad se viene incluyendo un problema de intervalo de confianza en prácticamente todas las convocatorias de los últimos años (López-Martín, Batanero, Díaz-Batanero y Gea, 2016).

Además, el intervalo de confianza es un tema básico en los cursos universitarios de inferencia en diversas titulaciones, entre otras la de Psicología, área de conocimiento que se

interesa en gran medida por la corrección del uso de la estadística en sus investigaciones (De la Fuente y Díaz-Batanero, 2004).

Es importante, entonces, asegurar que el estudiante alcanza un conocimiento suficiente del tema. Sin embargo, la investigación didáctica relacionada con este tema se ha concentrado en muestras de profesionales o estudiantes universitarios en carreras de ingeniería, teniéndose poca información sobre el grado de comprensión de los estudiantes españoles que se enfrentan a este tema en las pruebas de acceso a la universidad o de los estudiantes de ciencias humanas como la Psicología. Además, prácticamente todos los estudios previos se realizan mediante cuestionarios con ítems de opción múltiple que tratan de evaluar el conocimiento de la definición del intervalo de confianza y de sus propiedades.

El objetivo de este trabajo es complementar dicha investigación previa con información acerca de la comprensión de estudiantes de Bachillerato y de Psicología, cuestión que se aborda mediante la resolución de un problema abierto que permite estudiar cómo los estudiantes aplican y relacionan varios conceptos vinculados con el intervalo de confianza.

Fundamentos

Significado y comprensión de los objetos matemáticos

Nuestro trabajo utiliza las ideas de *significado* y *comprensión* que se deducen del enfoque ontosemiótico del conocimiento y la instrucción matemática (descritas, por ejemplo, en Godino 2002; Godino, Batanero y Font, 2007; 2019; y Godino, Batanero y Roa 2005). En este marco teórico se parte de la situación-problema, que se interpreta en sentido amplio como cualquier tarea, pregunta o cuestión que requiera una actividad de matematización. El significado de los objetos matemáticos surge de las prácticas realizadas en la resolución de problemas, pues son las situaciones en las que se debe utilizar un cierto objeto matemático las que le dotan de significado.

Una práctica es toda acción (práctica operatoria) o expresión (práctica discursiva) realizada por un sujeto para obtener la solución de un problema, generalizarla o comunicarla a otras personas. Además, en este marco teórico se diferencia entre prácticas personales o institucionales, es decir, las que son privativas de un sujeto y las realizadas por una institución de personas interesadas en resolver un mismo tipo de problemas. En base a ello, el significado

institucional o personal de un objeto se concibe como el sistema de prácticas institucionales o personales asociadas al mismo (Godino, Batanero y Font, 2007).

En el enfoque ontosemiótico se concibe la comprensión mostrada por el alumno como la coincidencia entre los significados institucional y personal sobre cada objeto matemático (Godino, 1996). Dicha comprensión tiene en cuenta diferentes aspectos del objeto cuya comprensión se quiere evaluar y se adquiere gradualmente. En este trabajo pretendemos evaluar la comprensión que muestran los estudiantes del objeto matemático “*intervalo de confianza*” a partir del análisis de las prácticas que realizan para resolver un problema abierto en el que deben construir un intervalo.

Significado frecuencial del intervalo de confianza

El problema de estimación por intervalos ha sido resuelto desde diferentes comunidades estadísticas que entiende la inferencia desde varios puntos de vista (frecuencial, bayesiano o remuestreo entre otros) y, por tanto, podemos considerar la existencia de significados institucionales diferentes de este objeto matemático (Olivo, 2008; Rivadulla, 1991). Dichos significados coexisten en la actualidad e implican diferentes prácticas matemáticas asociadas.

Todos estos enfoques coinciden en dar la solución al problema de estimar un parámetro θ de una población utilizando los datos de una muestra y un estimador $\hat{\theta}$ calculado a partir de ella, dando además una medida de la imprecisión de la estimación causada por la variabilidad del muestreo, proporcionando un rango de valores para el parámetro (Morey et al., 2016).

Destacamos los siguientes tres métodos de construcción e interpretación del intervalo de confianza. En este trabajo nos restringimos al *método frecuencial* que fue iniciado por Neyman (1937). Desde esta perspectiva se entiende que el parámetro poblacional θ es un valor constante y desconocido, y que el estimador $\hat{\theta}$ es una variable aleatoria (que cambia de una muestra a otra). En el *método bayesiano* se supone que el parámetro θ es una variable aleatoria con una distribución inicial de probabilidad y se utiliza la muestra seleccionada para obtener la distribución final del parámetro calculando, a partir de ella, el intervalo de confianza (llamado *intervalo de credibilidad*). En el *método de remuestreo* se parte de una muestra inicial de la que se extrae, por remuestreo, otras muchas muestras, a partir de las cuales se calcula una distribución de remuestreo del estimador $\hat{\theta}$ en las diferentes muestras, la cual se usa para obtener el intervalo de confianza (llamado, en este caso, *intervalo de remuestreo*).

La enseñanza actual en España se concentra en el método frecuencial, según el cual el valor de estimador $\hat{\theta}$ en la muestra se utiliza para determinar el intervalo de confianza. Esto significa que, para el caso de la media poblacional, se usa como estimador la media muestral. Se pretende obtener un intervalo con los siguientes extremos (EI y ES), asociados a un cierto *nivel de confianza* $(1 - \alpha)$, donde α es cualquier número entre 0 y 1, escogido anticipadamente y denominado *coeficiente de confianza* (Olivo, 1997):

$$EI = \hat{\theta} - k_1 S_{\hat{\theta}} \quad \text{y} \quad ES = \hat{\theta} + k_2 S_{\hat{\theta}}.$$

El coeficiente de confianza determina las constantes k_1 y k_2 que se necesitan en el cálculo del intervalo. Si la distribución muestral es simétrica alrededor del estadístico $\hat{\theta}$, estas constantes son iguales y se determinan obteniendo los valores centrales de dicha distribución que incluyen una probabilidad $1 - \alpha$.

Mayo (1981) recuerda que el nivel de confianza $1 - \alpha$ se refiere a los extremos del intervalo, que son variables aleatorias, y no representa la probabilidad de que el verdadero valor del parámetro θ esté situado dentro del intervalo. Es decir, se refiere al porcentaje de intervalos construidos a partir de diferentes muestras del mismo tamaño de la población que cubrirán el parámetro. Morey et al. (2016) sugieren la necesidad de diferenciar entre el procedimiento que genera el intervalo y el intervalo en sí mismo. Tenemos una confianza del $100(1 - \alpha)\%$ en que el procedimiento generará intervalos que, en dicho porcentaje, cubrirán el valor real del parámetro; pero cada intervalo en sí mismo no es aleatorio, sino fijo, y puede o no cubrir el parámetro. Los autores denominan *falacia fundamental de la confianza* al hecho de interpretar dicha confianza como probabilidad de que el parámetro esté contenido en el intervalo.

Antecedentes

Son pocas las investigaciones sobre comprensión del intervalo de confianza, y la mayor parte se han hecho con licenciados, con investigadores o con estudiantes de ingeniería. Así, Cumming, William y Fidler (2004) seleccionaron a 263 investigadores que habían publicado artículos de investigación en revistas internacionales en los cuales se utilizaba la estimación por intervalos. Tras proporcionarles un intervalo de confianza para la media de una población, construido a un nivel de confianza del 95%, se interrogó a dichos investigadores sobre los valores que ellos entendían que debían ser factibles para la media muestral si se repitiese el mismo experimento en un gran número de ocasiones. El 74% de los investigadores respondieron

que la media volvería a caer en el intervalo proporcionado en el 95% de las ocasiones (probabilidad de replicación).

Behar (2001) realizó un estudio con 47 profesionales de estadística y 297 estudiantes de ingeniería e identificó un error de interpretación, pues el 29% de los expertos y el 50% de los estudiantes afirmaba que el nivel de confianza indicaba el porcentaje de datos poblacionales que caen dentro del intervalo de confianza. Por otro lado, el 40% de los expertos y poco más de la mitad de los estudiantes indicaron que el nivel de confianza es la probabilidad de que el intervalo contenga al verdadero valor de la media poblacional. Otros participantes en el estudio creían erróneamente que cuando el nivel de confianza es alto, dejando invariantes los demás factores, se obtenían intervalos más estrechos, por lo que mostraban ausencia de comprensión de la relación entre la anchura del intervalo y el tamaño de la muestra.

En una serie sucesiva de trabajos (Olivo, 2008; Olivo, Díaz y Batanero, 2007; Olivo et al., 2008), se analizó la comprensión de alumnos de ingeniería mediante un amplio cuestionario (que puede consultarse en Olivo, 2008) en el que se recogían preguntas sobre diversos contenidos relacionados con este tema. Dicho cuestionario fue propuesto a una muestra de 48 estudiantes, y sus resultados fueron publicados en Olivo y Batanero (2007). Muchas de las cuestiones fueron correctamente respondidas por la mayor parte de participantes en el estudio. Sin embargo, también se detectaron muchas dificultades, además de los siguientes errores conceptuales:

- El 18% afirmaba que una variación en el nivel de confianza no afectaría a la anchura del intervalo.
- El 21% relacionaba el intervalo con posibles valores de la variable estudiada, en vez de relacionarlo con su media poblacional.
- El 27% de los participantes interpretaba el nivel de confianza con la probabilidad de que la media cayese dentro de él.

Respecto al conocimiento procedimental, en el trabajo de Olivo, Batanero y Díaz (2008) el 41,3% de los estudiantes de ingeniería tuvo errores al obtener el valor crítico para construir el intervalo, y el 18,7% construyó un intervalo incorrecto para la comparación de dos varianzas al intercambiar los grados de libertad de numerador y denominador. Igualmente, el 34,5% obtuvo unos grados de libertad incorrectos para determinar el intervalo de confianza de una

media, por lo que los autores sugirieron que no se comprendía el concepto de grados de libertad. Otro 11,5% confundió las desviaciones típicas de la población y de la distribución muestral.

Fidler y Cumming (2005) analizaron la interpretación del intervalo de confianza de 180 estudiantes de Psicología, los cuales, en su mayoría, apreciaban los intervalos de confianza desde una perspectiva descriptiva en lugar de valorar su carácter inferencial: un 38% imaginaba el intervalo como un conjunto de valores razonables para la media de la muestra, y un 19% confundía las nociones de rango (o recorrido de la variable aleatoria) e intervalo de confianza. Además, un 20% manifestó que la anchura del intervalo aumentaría cuando creciese el tamaño muestral, un 29% consideró que, en tal caso, la anchura no variaría y, finalmente, un 36% no fue capaz de describir si existía o no relación entre la anchura y el tamaño de la muestra.

En Roldán López de Hierro, Batanero y Álvarez-Arroyo (en prensa) analizamos la comprensión conceptual del intervalo de confianza de 58 estudiantes de Bachillerato y 37 de ingeniería a través de sus respuestas a seis cuestiones de opción múltiple. Los resultados muestran la presencia de todas las dificultades descritas en las investigaciones que acabamos de resumir, siendo mejor los resultados de los estudiantes de ingeniería.

Con objeto de ahondar en el análisis de la comprensión de los estudiantes en el tema, en este nuevo trabajo proponemos un problema abierto a la misma muestra de estudiantes de Bachillerato que participó en el trabajo de Roldán López de Hierro et al. (en prensa) y a otra muestra de estudiantes de Psicología, eligiendo este grupo por ser estudiantes de Ciencias Sociales, al igual que los estudiantes de Bachillerato de la muestra. La resolución del problema abierto nos permitirá analizar las prácticas matemáticas realizadas por los estudiantes para, a partir de ella, identificar la comprensión lograda de los objetos matemáticos que intervienen en el proceso de resolución.

Metodología

La muestra estuvo formada por dos tipos de estudiantes. En primer lugar, participaron 58 estudiantes de segundo curso de Bachillerato de Ciencias Sociales de dos institutos diferentes de la ciudad de Granada, el primero situado en el centro de la ciudad (24 estudiantes) y el segundo en la periferia (34 estudiantes). Estos estudiantes completaron el problema propuesto dentro de la clase de matemáticas, una vez estudiado el tema de los intervalos de confianza y durante el último trimestre del curso, cuando se preparaban para las pruebas de acceso a la

universidad, por lo que se estaban ejercitando en la resolución de problemas de inferencia, incluido el intervalo de confianza.

El segundo grupo estuvo constituido por 57 estudiantes de primer curso de Psicología de la Universidad de Huelva, que cursaban la asignatura Análisis de datos II, centrada en la inferencia estadística, durante el segundo semestre del curso. Los estudiantes habían estudiado otra asignatura denominada Análisis de datos I en el primer semestre, donde estudiaron estadística descriptiva y cálculo de probabilidades. Completaron el problema dentro de una evaluación parcial del curso, una vez que lo habían estudiado.

En consecuencia, la muestra es intencional, por lo que no pretendemos generalizar las conclusiones a una población más amplia, dado que el estudio es exploratorio. No obstante, pensamos que el análisis de las prácticas matemáticas que estos estudiantes realizan en la resolución del problema propuesto nos permite comprender mejor la complejidad del significado institucional del intervalo de confianza, así como de algunos significados personales que les atribuyen los estudiantes.

A los dos grupos de estudiantes se les propuso el siguiente problema abierto, que se tomó de la prueba de acceso a la universidad en Andalucía del año 2018. El motivo para elegir un problema de estas pruebas fue asegurarse de que era familiar para ambos tipos de estudiantes: para los de Bachillerato, porque se estaban preparando en ese momento para tales pruebas de acceso y resolvían problemas similares; para los de Psicología, porque habían realizado las pruebas de acceso el curso anterior.

Problema. La media muestral de 100 observaciones en una prueba de matemáticas es de 75 puntos. Asumiendo que las calificaciones siguen una distribución normal y que $\sigma = 7$, encuentre el intervalo de confianza al 95% para la media de la población.

El problema se puede resolver recordando y aplicando la fórmula de cálculo del intervalo de confianza para la media de una población (con desviación típica conocida), a partir de los datos del problema, que son el tamaño de muestra ($n = 100$), la media muestral $\bar{x} = 75$ y la desviación típica de la población $\sigma = 7$.

Fijado un *coeficiente de confianza* $1 - \alpha$ (en el problema, situado en el 95%), la estimación por intervalo de la media poblacional μ consiste en determinar un intervalo, centrado en la media muestral \bar{x} , de manera que una cierta proporción (aproximadamente igual al coeficiente de confianza) de intervalos construidos a partir de muestras diferentes, pero de igual

tamaño de la población, contenga a la media de la población. Se utiliza la media muestral como estimador de la media poblacional debido a las siguientes propiedades: es un estimador insesgado, de mínima varianza y consistente (se acerca al valor estimado al aumentar el tamaño de la muestra; Silvey, 2017).

Es muy importante recordar que no es posible saber si el verdadero valor de la media poblacional pertenece, o no, al intervalo de confianza una vez construido éste porque, desde el principio, asumimos que μ es desconocida. El nivel de confianza se refiere al método de construcción de los intervalos de forma que, si construimos un gran número de ellos a partir de muestras aleatorias independientes, la teoría estadística indica que al menos el $100(1 - \alpha)\%$ de los mismos contendrá al verdadero valor de la media poblacional (Roldán López de Hierro, 2019).

La determinación de la fórmula para obtener dicho intervalo está basada en la distribución muestral de la media de la muestra que es una variable aleatoria, ya que varía de una muestra a otra, y en las condiciones del problema sigue una distribución normal $N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$, es decir, su valor esperado es la media de la población y su desviación típica igual a la de la población dividida por la raíz cuadrada del tamaño n de la muestra. Esta desviación típica de la distribución muestral de la media $\sigma_{\bar{x}}$ se denomina *error típico*. Aplicando el proceso de tipificación, la variable aleatoria $Z = (\bar{X} - \mu)/\sigma_{\bar{x}}$ sigue una distribución normal estándar $N(0,1)$. Si denotamos por $z_{\alpha/2}$ al único valor real (positivo) que satisface la igualdad $P(Z > z_{\alpha/2}) = \alpha/2$ (al que llamaremos *nivel de riesgo* o *valor crítico* para el nivel de confianza $1 - \alpha$), entonces los extremos del intervalo de confianza, para un coeficiente de confianza $1 - \alpha$ se obtienen mediante la siguiente fórmula:

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

La semi-amplitud del intervalo de estimación, es decir, el número $z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$, se conoce como *error máximo de estimación*. Sustituyendo los datos del problema en esta fórmula, se obtiene el intervalo pedido:

$$\left[75 \pm 1.96 \frac{7}{\sqrt{100}} \right] = [73.628, 76.372]$$

Para analizar más profundamente esta solución, en el Cuadro 1 mostramos un análisis semiótico de la misma en base al marco teórico descrito, descomponiéndola en prácticas

matemáticas elementales y, para cada una de ellas, identificando los objetos y procesos matemáticos involucrados. Dicho análisis pone de manifiesto la gran cantidad de objetos matemáticos que se deben recordar y aplicar en la solución y, por tanto, la complejidad semiótica de la solución del problema.

Cuadro 1. Análisis semiótico de la solución al problema.

Práctica matemática	Objetos y procesos involucrados
PM1. Recordar o deducir la fórmula de cálculo del intervalo: $\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$	<ul style="list-style-type: none"> • Recordar los conceptos de intervalo y de extremos inferior y superior del mismo. • Diferenciar la media de la muestra (\bar{x}) y la media de la población (μ) y recordar que el intervalo está centrado en la media muestral, pues ésta es un estimador insesgado y de mínima varianza de μ. • Recordar que la amplitud del intervalo viene dada por el valor crítico multiplicado por la varianza de la distribución muestral. • Recordar que la desviación típica de la distribución muestral es $\frac{\sigma}{\sqrt{n}}$. • Recordar que el valor crítico se obtiene a partir del nivel de confianza y la distribución normal estándar.
PM2. Determinar el nivel de riesgo $\alpha/2$ a partir del nivel de confianza $1 - \alpha$	<ul style="list-style-type: none"> • Recordar los conceptos de nivel de confianza y de nivel de riesgo, y la relación entre ellos. • Determinar el nivel de riesgo $\alpha/2$ a partir del nivel de confianza.
PM3. Lectura de la tabla para determinar el valor crítico	<ul style="list-style-type: none"> • Recordar que, en este caso, se puede aplicar la distribución normal $N(0,1)$, previa tipificación de los datos. • Lectura de la tabla de la distribución normal $N(0,1)$, con posible necesidad de interpolación.
PM4. Determinar el error típico σ/\sqrt{n} y el error máximo $z_{\alpha/2} \sigma/\sqrt{n}$	<ul style="list-style-type: none"> • Recordar las fórmulas y los conceptos de error típico y de error máximo de estimación, y la forma en que intervienen en el cálculo del intervalo de confianza.
PM5. Sustitución de datos en la fórmula del intervalo	<ul style="list-style-type: none"> • Identificar los datos necesarios en la fórmula del intervalo a partir del enunciado del problema, recordando los conceptos requeridos.
PM6. Cálculos	<ul style="list-style-type: none"> • Realizar los cálculos requeridos para obtener los extremos del intervalo.
PM7. Redondeo	<ul style="list-style-type: none"> • Recordar los criterios de redondeo y aplicarlos correctamente, si es necesario.

Fuente: Elaboración Propia

Resultados

A continuación, presentamos los resultados obtenidos en nuestro estudio, incluyendo el análisis de algunos ejemplos de soluciones de los estudiantes. Para diferenciar el grupo denotaremos por P1, P2, etc., a las respuestas de los estudiantes de Psicología, y por B1, B2, etc., las de Bachillerato. En primer lugar, se clasificaron las soluciones de los estudiantes en “correctas”, “parcialmente correctas” e “incorrectas”.

Entendemos que una resolución “correcta” es aquella que lleva a una expresión final correcta del intervalo, con o sin redondeo. Un ejemplo de ello, dado por un estudiante de Psicología, se presenta en la Figura 1. En ella observamos que el estudiante sigue todos los pasos descritos en nuestro análisis de la solución del problema (Cuadro 1). Incluso explícitamente nombra muchos de los conceptos que va utilizando, como nivel de confianza o de riesgo, error típico y máximo, e intervalo de confianza, utilizando también una simbología correcta.

Figura 1. Resolución correcta del problema del estudiante de Psicología P1.

The image shows a handwritten solution for a confidence interval problem. The student starts with a confidence level of 0.95, which leads to a risk level of 0.05. They then calculate the critical value $z = 1.96$ by dividing 0.05 by 2. Next, they find the standard error by dividing the standard deviation (6) by the square root of the sample size (100), resulting in 0.6. The maximum error is calculated as $1.96 \cdot 0.6 = 1.176$. Finally, the confidence interval is determined by adding and subtracting the maximum error from the sample mean (75), resulting in the interval $(73.824, 76.176)$.

$$\begin{aligned} \text{N. confianza} &= 0.95 \rightarrow \text{N. riesgo} = 0.05 \rightarrow \text{no me} \\ \frac{\alpha}{2} &= \frac{0.05}{2} = 0.025 \rightarrow z = 1.96 & \text{Solución: intervalo de confianza} &= (73.824 - 76.176) \\ \text{Error típico} &= \frac{6}{\sqrt{n}} = \frac{6}{\sqrt{100}} = 0.6 \\ \text{Error máx.} &= 1.96 \cdot 0.6 = 1.176 \\ \text{Intervalo} &= 75 \begin{cases} + 1.176 = 76.176 \\ - 1.176 = 73.824 \end{cases} \end{aligned}$$

Fuente: Datos de la Investigación

Otro ejemplo, esta vez dado por un estudiante de Bachillerato, se presenta en la Figura 2 que, siendo también correcta y utilizando la simbología adecuada, no describe los conceptos que está utilizando, por lo que no queda tan claro que comprenda realmente todos estos conceptos. Sin embargo, este estudiante, al contrario que en el ejemplo anterior, realiza una gráfica de la distribución normal, como ayuda en la determinación del percentil $z_{\alpha/2}$ a partir del nivel de riesgo.

Figura 2. Resolución correcta del problema del estudiante de Bachillerato B1.

$$\begin{array}{l}
 n = 100 \\
 \bar{x} = 75 \\
 s = 7 \\
 E < 5?
 \end{array}
 \quad
 \begin{array}{l}
 1 - \alpha = 0.95 \\
 \alpha = 0.05 \\
 \frac{\alpha}{2} = 0.025
 \end{array}$$


$$\mu = \left(75 - 1.96 \frac{7}{\sqrt{100}} ; 75 + 1.96 \frac{7}{\sqrt{100}} \right)$$

$$(73.628 ; 76.372) //$$

Fuente: Datos de la investigación

Entendemos por resolución “parcialmente correcta” aquella en la que el estudiante llega a escribir la expresión del intervalo de confianza, sustituyendo correctamente los datos requeridos y, sin embargo, por una u otra razón (posiblemente un fallo en las operaciones), no llega a la expresión final del intervalo de confianza. Y una solución “incorrecta” es aquella en que la expresión del intervalo es incorrecta o no se deducen correctamente los datos necesarios para el cálculo a partir del enunciado del problema.

En la Tabla 1 se presentan los resultados sobre la corrección de la respuesta en los dos grupos. Observamos una gran diferencia en la proporción de soluciones correctas entre ambos grupos de estudiantes, siendo mucho mejor los resultados en los estudiantes de Psicología, con el 80% de las soluciones correctas y sin soluciones parcialmente correctas. Aun así, se obtiene todavía un porcentaje de estudiantes con soluciones incorrectas, aunque prácticamente ningún estudiante deja la respuesta en blanco.

Respecto al grupo de estudiantes de Bachillerato, entre soluciones correctas y parcialmente correctas obtenemos un 40%, y es notable que un 30% de los mismos deje la solución en blanco. Estos resultados muestran una dificultad excesiva del tema para este grupo, incluso después de prepararse para las pruebas de acceso donde con frecuencia se les propone un problema de este tipo. El gran porcentaje de respuestas incorrectas sugiere que es un mejor método tratar de que el estudiante explicita todos los pasos y conceptos utilizados en la construcción del intervalo (Figura 1) que simplemente enseñarle a aplicar una fórmula (Figura 2). La razón es que se puede recordar mal la fórmula o no ser capaz de deducir los datos necesitados para aplicarla.

Tabla 1.

Distribución de estudiantes según la resolución del problema.

Tipo de resolución	Bachillerato		Psicología	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
Resolución correcta	17	29,3	46	80,7
Resolución parcialmente correcta	6	10,4		
Resolución incorrecta	17	29,3	10	17,5
En blanco	18	31,0	1	1,8

Fuente: Datos de la Investigación

Un análisis más detallado del proceso de resolución del problema seguido por los estudiantes se muestra en la Tabla 2, en la que se presentan las frecuencias de estudiantes de cada uno de los dos grupos que llegan a completar las prácticas matemáticas analizadas en la Tabla 1, que son pasos necesarios en la resolución. La mayoría de los estudiantes de Psicología y alrededor del 40% de los de Bachillerato recuerdan y escriben la expresión del intervalo (PM1), y por tanto llegarán a soluciones correctas o parcialmente correctas.

También la mayoría de estudiantes de Psicología es capaz de determinar el nivel de riesgo $\alpha/2$ a partir del nivel de confianza $1 - \alpha$, mientras que apenas la mitad del alumnado de Bachillerato lo consigue. Consideramos muy preocupante esto último, pues dicho valor es el percentil que ha de usarse posteriormente para determinar el valor crítico, y pone de manifiesto que los estudiantes de Bachillerato de nuestra muestra desconocen el significado más esencial del proceso que van a tratar de desarrollar.

Entre los que determinan correctamente el valor de $\alpha/2$, la mayoría determina el valor crítico $z_{\alpha/2}$, leyendo correctamente la tabla de la distribución normal, con lo que muestran conocimiento de dicha distribución y del significado de valor crítico. Además, unos pocos alumnos escriben el valor crítico correcto sin haber explicitado previamente el nivel de riesgo $\alpha/2$ ni haber hecho una gráfica o esquema adecuado de la distribución normal que les ayude en dicho cálculo, por lo que podemos pensar que simplemente recuerdan que el valor crítico en la distribución normal para el nivel de confianza 0,95 es igual a 1,96.

Tabla 2.

Frecuencia de estudiantes que realizan de los pasos de una correcta resolución del problema.

Descripción del proceso seguido	Bachillerato		Psicología	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
PM1. Recuerda o deduce la expresión del intervalo	23	39,7	46	80,7
PM2. Determina el nivel de riesgo $\alpha/2$, dado el coeficiente de confianza $1 - \alpha$	30	51,7	52	91,2
PM3. Determina el valor crítico $z_{\alpha/2}$, leyendo correctamente la tabla de la distribución normal	26	44,8	43	75,4
PM4. Determina el error típico σ/\sqrt{n} y el error máximo $z_{\alpha/2} \sigma/\sqrt{n}$			39	68,4
PM5. Escribe correctamente la expresión del intervalo $\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$ sustituyendo los datos	23	39,7	4	7
PM6. Realiza correctamente los cálculos requeridos	17	29,3	46	80,7
PM7. Da la expresión final de los extremos:				
- Sin redondeo	14	24,1	35	61,4
- Redondeando	3	5,2	11	19,3

Fuente: Datos de la Investigación

Únicamente los estudiantes de Psicología describen y calculan explícitamente el error típico y el error máximo (PM4) como paso previo en la construcción del intervalo siendo, además, la mayoría de los que obtuvieron la solución correcta. Casi todos los estudiantes de Bachillerato que llegan a determinar correctamente el valor crítico y solamente cuatro de Psicología escriben directa y correctamente la fórmula del intervalo de confianza, sustituyendo correctamente los datos requeridos. Finalmente se realizan los cálculos correctamente (todas las soluciones correctas de ambos grupos) y se expresan los valores obtenidos para los extremos del intervalo, redondeando en algunos casos a 1 o 2 decimales. Por otro lado, 13 alumnos de Bachillerato y 3 de Psicología se apoyan en una representación gráfica correcta de la distribución normal para la obtención del valor crítico (como la de la Figura 2), llegando 10 de Bachillerato y esos 3 de Psicología a la solución correcta del problema, lo cual parece indicar que el uso de dicha representación gráfica facilita una visión global adecuada del proceso que están llevando a cabo y una mejor comprensión de esa práctica.

Para completar el análisis mostramos en la Tabla 3 los errores identificados en las respuestas parcialmente correctas o incorrectas, con el porcentaje de aparición respecto al total de alumnos en cada grupo. Algunos alumnos tienen más de un error en el proceso de solución. Muchos de estos errores se habían descrito en la investigación previa, por lo que era previsible su aparición en nuestro estudio.

Tabla 3.

Errores más frecuentes en el proceso de resolución

Tipo de error	Bachillerato		Psicología	
	Frecuencia	%	Frecuencia	%
<i>Errores encontrados en trabajos previos</i>				
Determina incorrectamente el valor crítico $z_{\alpha/2}$	7	12,1	9	15,8
Error en operaciones o uso de expresiones algebraicas	5	8,6	1	1,8
<i>Errores identificados en el estudio</i>				
Suma y resta a la media el valor crítico antes de multiplicar por el error típico	4	6,9		
Suma y resta a la media el valor crítico multiplicado por el tamaño de la muestra	1	1,7		
Determina incorrectamente nivel de riesgo $\alpha/2$	1	1,7	3	5,3
Redondeos incorrectos	9	15,5		
Toma un valor crítico negativo $-1,96$			9	15,8
Intercambia los extremos del intervalo			1	1,8

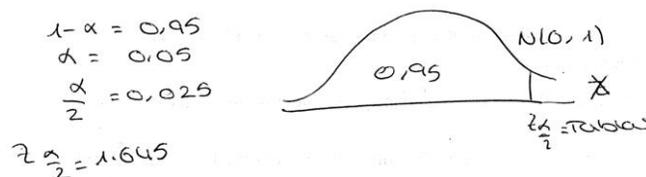
Fuente: Datos de la Investigación

Errores descritos en la investigación previa

El error más frecuente (12,1 % de estudiantes de Bachillerato y 15,8% de Psicología) es determinar incorrectamente el valor crítico $z_{\alpha/2}$ debido a lectura incorrecta de la tabla, a pesar de que el valor crítico correspondiente al 95% de confianza que enunciaba el problema se utiliza con mucha frecuencia en clase, lo que podría facilitar a los estudiantes el recordar que su valor es de 1,96. Este error también aparece en el trabajo de Olivo, Batanero y Díaz (2008) en el 41,3% de los estudiantes, por lo que nuestros resultados son mejores que los del citado trabajo.

En unos casos, como en el mostrado en la Figura 3, el estudiante representa incorrectamente el área correspondiente al nivel de confianza en la curva normal, confundiéndolo con la correspondiente a la cola izquierda, es decir, no comprende que se debe tomar el 95% central de los valores de la distribución normal (percentil del 97,5%) y, en su lugar, toma el percentil del 95%. Ello implica la falta de comprensión del concepto de coeficiente de confianza. Otros estudiantes no habían confeccionado una representación gráfica, por lo que no podían apoyarse en la misma para hacer una lectura adecuada de la tabla de la distribución normal.

Figura 3. Cálculo incorrecto del valor crítico por el estudiante de Bachillerato B2.



Fuente: Datos de la Investigación

En el ejemplo mostrado en la Figura 4, el estudiante confunde el valor crítico (que es un percentil de la curva normal) con el nivel de riesgo (que es una probabilidad), por lo que no usa la tabla de la distribución normal para deducir el valor crítico y obtiene un intervalo incorrecto.

Figura 4. Confusión de valor crítico y nivel de riesgo por el estudiante de Bachillerato B3.

$$\text{Fórmula} = \left(\bar{x} - z_{\alpha/2} \cdot \frac{s}{\sqrt{n}}, \bar{x} + z_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \right)$$

$$\left(75 - 0,025 \cdot \frac{7}{\sqrt{100}}, 75 + 0,025 \cdot \frac{7}{\sqrt{100}} \right) = \left(75 - 0,025 \cdot 0,7, 75 + 0,025 \cdot 0,7 \right)$$

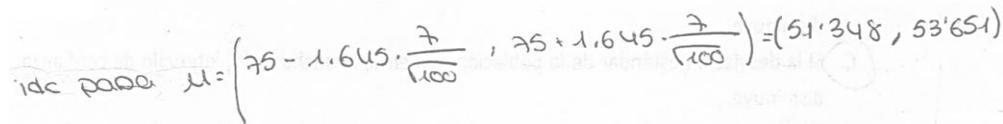
Fuente: Datos de la Investigación

En ocasiones se producen errores al realizar las operaciones; en unos casos se trata de errores aritméticos y, en otros, como en el ejemplo de la Figura 3, se trata de estudiantes que no dominan el trabajo con expresiones algebraicas (es más usual encontrar estudiantes cuyos conocimientos algebraicos previos son más deficientes en el Bachillerato de Ciencias Sociales que en el Científico y Tecnológico).

Nuevos errores identificados en nuestro estudio

Los nuevos errores identificados que más llaman la atención son aquellos que, una vez producidos, indican objetivamente que una parte del razonamiento es incorrecto. En este sentido, es especialmente llamativo que el alumnado haya sumado y restado a la media muestral el valor crítico, antes de multiplicar por el error típico (desviación típica de la distribución muestral), dando lugar al intervalo erróneo generado por la expresión $(\bar{x} \pm z_{\alpha/2}) \cdot \frac{\sigma}{\sqrt{n}}$. Al evaluar esta expresión algebraica se obtiene el intervalo $[51,12; 53,87]$, el cual es claramente incorrecto pues no contiene a la media muestral $\bar{x} = 75$. Un ejemplo de este error puede observarse en la Figura 5, donde el estudiante escribe la fórmula correcta del intervalo, pero no tiene suficiente dominio de la expresión algebraica y se produce el error descrito. Obsérvese que el estudiante también había determinado incorrectamente el valor crítico, obteniendo 1,645.

Figura 5. Respuesta del estudiante de Bachillerato B4.

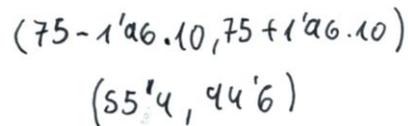


ide para $\mu = \left(75 - 1.645 \cdot \frac{7}{100}, 75 + 1.645 \cdot \frac{7}{100} \right) = (51.348, 53.651)$

Fuente: Datos de la Investigación

Otro tipo de fórmula errónea, como en el ejemplo mostrado en la Figura 6, es que en lugar de multiplicar el valor crítico por el error típico se multiplica por el tamaño de la muestra:

Figura 6. Respuesta del estudiante de Bachillerato B5, donde suma y resta a la media muestral el error típico multiplicado por el tamaño de la muestra.



$(75 - 1'96 \cdot 10, 75 + 1'96 \cdot 10)$
 $(55'4, 44'6)$

Fuente: Datos de la investigación

Cuatro estudiantes determinan erróneamente el nivel de riesgo al usar directamente α en lugar de $\alpha/2$; son estudiantes que muestran no comprender este concepto y arrastran el error al determinar un valor crítico incorrecto, que da lugar a extremos erróneos en el intervalo de confianza.

Respecto a la cuestión del redondeo de los extremos del intervalo de confianza, lo que se observa es que el alumnado trunca el número que obtiene en la calculadora sin más

consideración, lo cual es preocupante si tenemos en cuenta que es un error de cursos muy anteriores (primero o segundo de la E.S.O.) y aquí ha sido cometido por el 15,5% de nuestros alumnos de segundo de Bachillerato.

Otros errores aislados son los siguientes: a) intercambiar los extremos del intervalo, lo que indica desconocimiento de este concepto y del orden numérico; b) utilizar un valor crítico correcto pero negativo, debido a que busca en la curva normal el percentil del 2,5% en lugar del percentil de 97,5%, y que conlleva como resultado que en ocasiones obtenga cambiados los extremos del intervalo, aunque otras los escriba en orden correcto; y c) calcular $\frac{z_{\alpha}}{2}$ en lugar de $z_{\alpha/2}$, lo que de nuevo muestra confusión en la idea de nivel de riesgo.

Discusión e implicaciones didácticas

El intervalo de confianza proporciona más información que los contrastes de hipótesis, permitiendo cambiar el razonamiento estadístico dicotómico (rechazar o no una hipótesis) por un razonamiento de estimación, que se enfoca en la magnitud del efecto de las variables investigadas (Cumming, 2013). Es, por tanto, importante ayudar al estudiante a adquirir este tipo de razonamiento. Por ello, nuestro análisis ofrece una información relevante a los profesores de matemáticas de Bachillerato y profesores de estadística en Psicología, que podrían no ser conscientes de la complejidad y dificultad que el tema tiene para sus alumnos.

Sería necesario hacer reflexionar a los estudiantes sobre los pasos que se siguen en la construcción del intervalo de confianza y pedirles describir los objetos matemáticos que intervienen en cada uno de estos pasos. Debido al poco tiempo disponible, en ocasiones se enseña a construir el intervalo de confianza únicamente memorizando y aplicando la fórmula, sin una reflexión profunda sobre estos pasos de la construcción.

Es evidente que hay que superar este problema pues el cálculo de intervalos de confianza es una cuestión ampliamente resuelta por la tecnología, y la construcción del mismo no debe ser el principal objetivo de la enseñanza. Recomendamos al profesorado destinar más tiempo a transmitir a los estudiantes una comprensión de los diferentes objetos matemáticos involucrados en esta construcción y de sus propiedades, entre ellas las siguientes (nos referimos sólo a la media, aunque la discusión se aplica a otros parámetros, como la proporción):

- Es necesario diferenciar claramente la muestra de la que se dispone, con su media y desviación típica muestral, de la población objeto de estudio, modelizada por una variable con media y desviación típica poblacionales.
- El intervalo de confianza permite obtener una estimación de la media poblacional (que es un parámetro), a partir de la media muestral.
- Aunque usamos únicamente una muestra, es posible construir infinidad de muestras aleatorias del mismo tamaño que ésta. La media de cada una de todas estas muestras es una variable aleatoria y su distribución es la distribución muestral. La media de esta distribución muestral coincide con la media de la población y la desviación típica (error típico) es la desviación típica de la población dividida por la raíz cuadrada del tamaño de la muestra.
- El intervalo de confianza proporciona una estimación aproximada de la media poblacional a partir de la media muestral, a la que se suma y resta una cantidad, por lo que conviene asegurarse que el intervalo contiene a la media muestral. Los extremos del intervalo se obtienen teniendo en cuenta el nivel de confianza y la distribución muestral de la media de la muestra.
- Es importante recordar que no se sabe si el intervalo de confianza concreto construido en un estudio contiene al verdadero valor del parámetro poblacional. Pero sí se puede conocer la proporción de intervalos, generados con muchas muestras aleatorias independientes de igual tamaño, que lo contienen y que viene dada por el coeficiente de confianza.
- Es importante apoyarse en las representaciones gráficas y simulaciones que permiten aumentar la comprensión del significado y propiedades del intervalo de confianza. Más aún, teniendo en cuenta que la mayor parte del alumnado que realiza una representación gráfica adecuada calcula sin dificultad el valor crítico correcto. Sería también importante corregir errores impropios de estudiantes de Bachillerato o universidad, como el redondeo, meras operaciones aritméticas o el uso incorrecto de expresiones algebraicas sencillas.

Finalmente resaltamos la necesidad de dedicar más tiempo al estudio del intervalo de confianza, pues será una noción de enorme importancia en los estudios posteriores del alumnado.

Agradecimientos

Proyectos TIN2017-89517-P y Grupos de Investigación FQM-268 y FQM-126 (Junta de Andalucía).

Referencias

- Behar, R. (2001). *Aportaciones para la mejora del proceso de enseñanza-aprendizaje de la estadística*. Tesis Doctoral. Universidad Politécnica de Cataluña.
- Coulson, M., Healey, M., Fidler, F., & Cumming, G. (2010). Confidence intervals permit, but do not guarantee, better inference than statistical significance testing. *Frontiers in Psychology, 1*(26), 1-9. <https://doi.org/10.3389/fpsyg.2010.00026>
- Cumming, G. (2013). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. Routledge.
- Cumming, G., Williams, J., & Fidler, F. (2004). Replication, and researchers' understanding of confidence intervals and standard error bars. *Understanding Statistics, 18*(3), 299-311. DOI: [10.1111/j.1467-9280.2007.01881.x](https://doi.org/10.1111/j.1467-9280.2007.01881.x)
- Fidler, F., & Cumming, G. (2005). Teaching confidence intervals: Problems and potential solutions. *Proceedings of the 55th International Statistics Institute Session CD-ROM*. Sidney, Australia: International Statistical Institute.
- De la Fuente, E. I. & Díaz-Batanero, C. (2004). Controversias en el uso de la inferencia en la investigación experimental. *Metodología de las Ciencias del Comportamiento, Volumen especial 2004*, 161-167.
- Godino, J. D. (1996). Mathematical concepts, their meaning, and understanding. En L. Puig y A. Gutierrez (Eds.), *Proceedings of XX Conference of the International Group for the Psychology of Mathematics Education* (v.2, pp. 417-425). Universidad de Valencia.
- Godino, J. D. (2002). Un enfoque ontológico y semiótico de la cognición matemática. *Recherches en Didactique des Mathématiques, 22* (2 y 3), 237-284.
- Godino, J. D., Batanero, C. & 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., Batanero, C. & Font, V. (2019). The onto-semiotic approach: Implications for the prescriptive character of didactics. *For the Learning of Mathematics, 39* (1), 38-43.
- Godino, J. D., Batanero, C., & Roa, R. (2005). An onto-semiotic analysis of combinatorial problems and the solving processes by university students. *Educational Studies in Mathematics, 60* (1), 3-36.
- López-Martín, M. M., Batanero, C., Díaz-Batanero, C. & Gea, M. (2016). La inferencia estadística en las Pruebas de Acceso a la Universidad en Andalucía, *Revista Paranaense de Educação Matemática, 5*(8), 33-59.
- Mayo, D. G. (1981). In defense of the Neyman-Pearson theory of statistics. *Philosophy of Science, 48*, 269-280. DOI: <http://dx.doi.org/10.1590/1980-4415v33n64a11>.

- Ministerio de Educación, Cultura y Deporte, MECD (2015). *Real Decreto 1105/2014, de 26 de diciembre, por el que se establece el currículo básico de la Educación Secundaria Obligatoria y del Bachillerato*. Madrid: Autor.
- Morey, R. D., Hoekstra, R., Rouder, J. N., Lee, M. D., & Wagenmakers, E. J. (2016). The fallacy of placing confidence in confidence intervals. *Psychonomic Bulletin & Review*, 23 (1), 103-123. DOI: <https://doi.org/10.3758/s13423-015-0947-8>.
- Neyman, J. (1937). Outline of a theory of statistical estimation based on the classical theory of probability. *Philosophical Transaction of the Royal Society of London, series A, Mathematical and Physical Sciences*, 236 (767), 33-380.
- Olivo, E. (2008). *Significados del intervalo de confianza en la enseñanza de la ingeniería en México*. Tesis Doctoral. Universidad de Granada.
- Olivo, E., & Batanero, C. (2007). Un estudio exploratorio de dificultades de comprensión del intervalo de confianza. *Unión*, 12, 37-51.
- Olivo, E., Batanero, C., & Díaz, C. (2008). Dificultades de comprensión del intervalo de confianza en estudiantes universitarios. *Educación Matemática*, 20 (3), 5-32.
- Rivadulla, A. (1991). *Probabilidad e inferencia científica*. Barcelona: Anthropos.
- Roldán López de Hierro, A. F. (2019). *Un análisis exploratorio de la comprensión del intervalo de confianza por estudiantes de bachillerato*. Tesis de Máster. Universidad de Granada.
- Roldán López de Hierro, A. F., Batanero, C., & Álvarez-Arroyo, R. (en prensa). Conflictos semióticos relacionados con el intervalo de confianza en estudiantes de Bachillerato e Ingeniería. *Educación Matemática Debate*.
- Silvey, S. D. (2017). *Statistical inference*. London: Routledge.
- Wilkinson, L. & Task Force on Statistical Inference (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54, 594-604. DOI: <https://doi.org/10.1037/0003-066X.54.8.594>
- Yaremko, R. M., Harari, H., Harrison, R. C. & Lynn, E. (2013). *Handbook of research and quantitative methods in psychology: For students and professionals*. Hilldale, NJ: Erlbaum.