

Razonamiento estadístico, razonamiento condicional y conocimientos previos en estadística, evaluación y covariación

Statistical reasoning, conditional reasoning, and prior knowledge in statistics, assessment and covariation

DOI: <https://doi.org/10.32870/dse.v0i32.1597>

Juan Gerardo Martínez-Borrayo*
Luis Alfredo Mayoral-Gutiérrez**

Resumen

La enseñanza de matemáticas debe fomentar el razonamiento lógico, especialmente en estadística, donde se destacan conceptos como pensamiento y razonamiento estadístico. Aun cuando la definición de Garfield y Chance (2000) es la más aceptada, siguen existiendo diversas definiciones de razonamiento estadístico, sin un consenso claro sobre su significado y evaluación. El objetivo del presente estudio fue explorar la relación y evaluación de este concepto junto con el razonamiento condicional evaluado con la tarea conocida de Watson (1968), en su versión abstracta y concreta, controlando además el conocimiento previo en estadística. La investigación tuvo un diseño cuasi experimental con medidas cuantitativas para las tres variables de interés; razonamiento estadístico, conocimiento previo en estadística y razonamiento condicional, las ejecuciones se evaluaron tomando como base la taxonomía SOLO de Biggs y Collins (1982, 1989). Los resultados revelaron un bajo desempeño en las tres variables de interés. Se encontró una correlación positiva entre razonamiento abstracto e intermedio, y una correlación negativa entre conocimientos en estadística y razonamiento abstracto.

Palabras clave: Estadística – razonamiento estadístico – taxonomía SOLO – razonamiento condicional – conocimientos previos.

Abstract

The teaching of mathematics teaching should encourage logical reasoning, especially in statistics, where concepts such as statistical thinking and reasoning stand out. Even though the definition by Garfield and Chance (2000) is the most accepted, there are still various definitions of statistical reasoning, with no clear consensus on its meaning and assessment. The aim of this study was to explore the relationship and assessment of this concept together with conditional reasoning evaluated with the well-known task

* Doctor en Neurociencias. Líneas de Investigación: procesos cognoscitivos y de aprendizaje; neuroeducación. Profesor-investigador, Departamento de Neurociencias, Universidad de Guadalajara. México. juan.gerardo.martinez.borrayo@gmail.com

** Doctor en Educación. Líneas de Investigación: procesos cognoscitivos y de aprendizaje. Profesor-investigador, Departamento de Estudios en Educación, Universidad de Guadalajara. México. amayoral@academicos.udg.mx

by Watson (1968), in its abstract and concrete version, controlling also prior knowledge in statistics. The research had a quasi-experimental design with quantitative measurements for the three variables of interest: statistical reasoning, prior knowledge in statistics and conditional reasoning. The performances were evaluated based on the SOLO taxonomy by Biggs and Collins (1982, 1989). The results revealed a low performance in the three variables of interest. A positive correlation was found between abstract and intermediate reasoning, as well as a negative correlation between knowledge in statistics and abstract reasoning.

Keywords: Statistics – statistical reasoning – SOLO taxonomy – conditional reasoning – prior knowledge.

Introducción

En la enseñanza de las matemáticas se pretende que los alumnos no solo sean capaces de “tener conocimiento de técnicas y conceptos matemáticos”, sino que también deben fomentar “el razonamiento lógico” (SEP, 2017); esta relación entre matemáticas y razonamiento es un aspecto subrayado por diversos autores (Stacey, 2006); y la estadística, como una parte de las matemáticas, también está muy interesada en esa relación.

Cuando se combina la enseñanza de las habilidades del pensamiento y de las destrezas matemáticas, en particular en estadística, es cuando se inicia a discutir sobre *pensamiento matemático* (Márquez, 2012), *pensamiento estadístico* (Ben-Zvi, Garfield, 2004), *literacidad estadística* (Ziegler, 2014) y *razonamiento estadístico* (Garfield, 2003).

En una revisión amplia se encontraron distintas definiciones de razonamiento estadístico (Chance *et al.*, 2000; Ben-Zvi, Garfield, 2004; Chan, Ismail, 2014; Chan *et al.*, 2015, 2016; Duarte, Cázarez, 2014; Garfield, 2003; Kinnear, 2013; Tempelaar, 2004; Wang *et al.*, 2009; Garfield, Chance, 2000; Lovett, 2001; delMas, 2002; Reading, 2002; Nicholson *et al.*, 2006; Garfield, Gal, 1999; delMas, 2004; Al-Labadi *et al.*, 2018; Evans, 2019; Burrill, 2020; Gómez-Blancarte, Chávez, 2022; Azmay *et al.*, 2023; HuserEAU *et al.*, 2024), de lo cual se constata que no existe consenso en el área sobre lo que significa el razonamiento estadístico, las implicaciones para su correcta evaluación y que difícilmente parten de investigaciones sobre lo que es el razonamiento humano. La mayoría de los autores asumen la definición de Garfield y Chance (2000: 101):

Statistical reasoning may be defined as the way people reason with statistical ideas and make sense of statistical information. This involves making interpretations based on sets of data, representations of data, or statistical summaries of data. Students need to be able to combine ideas about data and chance, which leads to making inferences and interpreting statistical results.¹

Pero esta definición no se acerca a lo que algunos autores sobre el razonamiento conciben como tal; delMas (2004) realizó una revisión de la investigación relevante sobre el razonamien-

¹ El razonamiento estadístico puede definirse como la forma en que las personas razonan con ideas estadísticas y dan sentido a la información estadística. Esto implica realizar interpretaciones basadas en conjuntos de datos, representaciones gráficas o resúmenes estadísticos. Los estudiantes deben ser capaces de combinar ideas sobre datos y probabilidad, lo que les permite realizar inferencias e interpretar resultados estadísticos.

to y usó esos hallazgos para identificar áreas del razonamiento estadístico que los estudiantes pueden encontrar más difíciles de aprender. Afirma que no existe consenso en la definición de razonamiento y suele usarse en lugar de otros conceptos como pensamiento, solución de problemas, toma de decisiones y pensamiento crítico. DelMas coincide con la definición que propone Galotti (1989: 333):

Mental activity that consists of transforming given information (called the set of premises) in order to reach conclusions. This activity must be focused on at least one goal (but maybe focused on more than one). The activity must not be inconsistent with systems of logic when all the premises are fully specified, although there may not always be an applicable system of logic to govern specific instances of reasoning...²

El razonamiento estadístico tiene que ver con la capacidad de *explicar* por qué se produjeron ciertos datos, y además ser capaz de *justificar* una conclusión con base en esos datos; en resumen, darles sentido a los resultados estadísticos. Explicar y justificar, son puntos clave en el área del razonamiento estadístico. Sin embargo, los estudiosos del razonamiento no dirían que esos son rasgos definitorios. La versión formal de lo que es el razonamiento está en concordancia con las reglas de la lógica, la teoría de la probabilidad y otras instancias que se configuran como normativas (Stein, 1996).

Taxonomía SOLO

Podemos colegir que si alguien llega a razonar siguiendo las normas de la lógica, estaría entonces en el nivel más alto de razonamiento. Todos aquellos que no logren ese nivel estarían por debajo en su habilidad para razonar. Sin embargo, dado que los investigadores parten del supuesto de que no hay una forma de definir el razonamiento, entonces han optado por recurrir a métodos que clasifican niveles de desempeño intelectual.

La más recurrida es la propuesta de Biggs y Collis (1982, 1989), la taxonomía SOLO (*Structure of the Observed Learning Outcome*). Este modelo neopiagetano propone cinco niveles en el desarrollo cognitivo: *prestructural*: en este nivel, los participantes no plantean ninguna conclusión o inferencia, o las que plantean no se refieren a aspectos clave de razonamiento de inferencias formales; *uniestructural*: los individuos se enfocan en algún aspecto relevante de la tarea planteada o en alguna etapa de la tarea, pero no son capaces de realizar alguna conexión o relación entre conceptos o procesos; *multiestructural*: los sujetos plantean una conclusión y/o hacen una generalización refiriéndose a dos o más aspectos clave, sin lograr integrar lo declarado; *relacional*: integra las partes entre sí, de tal manera que el todo tiene una estructura coherente

² La actividad mental consiste en transformar la información dada (llamada conjunto de premisas) con el objetivo de llegar a conclusiones. Esta actividad debe estar enfocada en al menos un objetivo (aunque pueden ser más de uno). Además, la actividad no debe ser inconsistente con los sistemas lógicos cuando todas las premisas están completamente especificadas, aunque no siempre pueda existir un sistema de lógica aplicable para casos específicos de razonamiento...

y con significado; y el nivel de *abstracción extendida*: las personas cumplen con las exigencias cognitivas del nivel relacional, pero además son capaces de transferir los conceptos y procesos fuera del contexto en que fueron aprendidos.

Desde el modelo del desarrollo cognitivo, los estudios en razonamiento estadístico nos dicen que, a nivel de educación básica, de 20 estudiantes de primero a quinto grado (4 alumnos por grado), solo uno llegó al nivel 4 o relacional (Jones *et al.*, 2000), o bien, que hacia el final de un entrenamiento en pensamiento estadístico de nueve sesiones (2 por semana) a 19 alumnos de segundo grado de primaria, 84% de ellos exhibían estar al menos en el nivel 2 o transicional, pero ninguno se encontraba en el nivel más bajo, es decir, en el idiosincrático (Jones *et al.*, 2001).

Al parecer, hay un desarrollo en la comprensión y el razonamiento de los datos en los estudiantes conforme avanzan en sus grados escolares; en el paso de primaria a secundaria, al estar todos ubicados en los niveles más bajos de la taxonomía sin haber ninguno en la parte alta de la clasificación, y al final cada vez más alumnos terminan ubicándose en las partes más altas de los niveles; Watson y Moritz (1998) confirman este hallazgo al explorar las respuestas individuales de 88 alumnos de tercero de primaria a tercer grado de secundaria.

Sin embargo, no todos los estudios concluyen que al terminar la primaria los alumnos estén preparados para razonar un poco mejor en estadística. Watson *et al.* (1995) clasificaron las respuestas de siete alumnos en total, seis de sexto grado de primaria y uno de primero de secundaria; encontraron que la mitad de los alumnos de primaria estaba en los niveles más bajos y la otra mitad en el nivel intermedio, mientras que el alumno de secundaria estuvo en el nivel más alto. O bien ninguno de los estudiantes de secundaria pudo llegar al nivel 4 o relacional, como lo confirma el estudio de Mooney (2002) con 12 estudiantes de secundaria.

Avanzar en los grados académicos no asegura que los estudiantes van a estar ubicados en los niveles más altos de la clasificación de Biggs y Collins ya que, aunque en una investigación con 10 estudiantes de primer año de preparatoria fueron capaces de llegar al nivel de abstracción extendida, ninguno de ellos obtuvo el mismo nivel en cuatro constructos que analizaron (descripción, organización y reducción, representación y análisis e interpretación de datos) (Chan *et al.*, 2016). Más aún, en una muestra de 23 estudiantes de un diplomado en estadística aplicada, ninguno estuvo en el nivel preestructural; 54% en el uniestructural; 41% en el multiestructural; 4.5% en el relacional y ninguno en el abstracto extendido (Li, Goos, 2013). En otro estudio se encontró que solo una estudiante de posgrado con curso de estadística se ubicó en el nivel relacional (Nor, Idris, 2010); y en uno más se encontró, incluso, que profesores de matemáticas de bachillerato, al ser evaluados en la comprensión y razonamiento estadístico se ubican en los niveles más bajos del modelo SOLO (Juárez, Inzunza, 2014).

En general, aunque hay algunos estudios que encuentran a los estudiantes de estadística en niveles altos de la taxonomía, la mayoría de ellos se ubican en los niveles bajos, o bien, en el mejor de los casos, en el nivel medio.

El razonamiento y su evaluación

El razonamiento ha sido estudiado en diversas áreas, pero la división más común es la que se hace entre razonamiento inductivo y deductivo. Dentro de esta última se suele hacer la división entre inferencia transitiva, el silogismo categorial y el razonamiento proposicional; esta última es la que utiliza operadores lógicos como la negación, conjunción, disyunción, la implicación, el bicondicional y otros más, pero el que sin duda ha recibido más atención es el condicional.

El condicional tiene la estructura del tipo “si ..., entonces...”, en donde la parte del *si* es el antecedente y la parte del *entonces* es el consecuente. Este tipo de razonamiento tiene reglas válidas y no válidas; las válidas son el *modus ponens* (si el antecedente es verdadero, el consecuente es verdadero) y el *modus tollens* (si el consecuente es falso, entonces el antecedente es falso); mientras que las inferencias inválidas son la afirmación del consecuente (si el consecuente es verdadero, entonces el antecedente es verdadero) y la negación del antecedente (si el antecedente es falso, entonces el consecuente es falso).

Este tipo de razonamiento ha sido ampliamente estudiado, sobre todo utilizando la tarea de selección de Wason (1966), también conocida como la tarea de las cuatro tarjetas. En esta tarea se presenta la siguiente condición: hay cuatro tarjetas que en cada uno de sus lados tienen letras o números; se les dice que si hay una E en un lado, entonces hay un 4 en el otro lado. Se le advierte al sujeto que la frase solo se refiere exclusivamente a las cuatro tarjetas, y que cada tarjeta tiene una letra en un lado y un número del otro lado. La tarea consiste en señalar únicamente aquellas tarjetas que es necesario volver para comprobar si la frase condicional es verdadera o falsa.

Desde entonces se han desarrollado otras formas de evaluar el razonamiento condicional, como son el *Jansson Conditional Reasoning Test*, el cual viene en dos versiones, la forma CF (con contenido concreto y familiar) y la forma SG (otra vez con contenido familiar, pero que al menos una parte de las premisas son contrarias a hechos observables) (Jansson, 1978). La prueba de Hadar y Henkin (1978), que consta de 32 ítems, cada uno de los cuales comienza con dos premisas, para dar las cuatro formas lógicas de *modus ponens*, *tollens*, afirmación del consecuente y negación del antecedente. La premisa condicional en cada par está redactada en una de las siguientes cuatro formas de negación: del antecedente, del consecuente, de ambos o de ninguno. La prueba de razonamiento de Deanna Kuhn (1977), basada en la teoría de Piaget de que en la fase de operaciones concretas se tiene la capacidad para trabajar con formas silogísticas simples, mientras que declaraciones condicionales más complejas requieren operaciones formales. Para probar esto, usaron conversaciones simples en un contexto lo más naturalista posible. Se les presentaban silogismos, 4 por sesión. O bien el *Cornell Conditional Reasoning Test*, forma X, que consiste en 72 preguntas, que evalúa 12 diferentes patrones de inferencia (Nolan, Brandon, 1984).

Con estas y otras pruebas de evaluación del razonamiento se ha observado que se mejora la respuesta si es presentada en términos menos abstractos (Johnson-Laird *et al.*, 1972). O cuan-

do del material a ser juzgado tiene contenidos significativos (Griggs, Cox, 1982). En este último texto, en el experimento 3 es que aparece la prueba conocida como “el problema de la edad para beber alcohol” el cual utilizamos en esta investigación.

Conocimientos previos en estadística

Los estudiantes, cuando llegan al aula, traen conocimientos previos que inciden tanto de manera positiva como negativa en los aprendizajes y pueden ser tanto implícitos como explícitos. Los conocimientos implícitos pueden ser esquemas (Ausubel *et al.*, 1976), o bien pueden ser conocimientos de sentido común que los sujetos generan por todo su bagaje de experiencias previas que ha tenido (Vosniadou *et al.*, 2008). Sin embargo, son los conocimientos previos explícitos los que abordamos en nuestra investigación. En relación con las estadísticas, se ha observado que hay una correlación entre esos conocimientos previos y las puntuaciones en diferentes pruebas de evaluación de los conocimientos y habilidades estadísticas. Por ejemplo, se ha encontrado que las actitudes hacia la estadística están significativamente relacionadas con variables cognitivas como habilidades matemáticas básicas, conocimientos previos en estadísticas y los cursos tomados anteriormente (Roberts, Saxe, 1982). También se ha constatado que, en las pruebas de evaluación de la ansiedad en el aprendizaje de las estadísticas, hay una relación con el nivel de la clase en la que está el estudiante, su logro como estudiante, su bagaje matemático y sus cursos previos de estadística (Sutarso, 1992). De igual manera, se ha investigado la forma en que las actitudes y creencias de los estudiantes pueden impedir (o ayudar) en el aprendizaje de las estadísticas y que puede afectar el nivel en que los estudiantes desarrollan habilidades de pensamiento estadístico y también su habilidad de aplicar lo aprendido fuera del salón de clases (Gal *et al.*, 1997).

Desarrollo

Planteamiento del problema

Nuestro objetivo en esta investigación es comenzar a vincular dos áreas de trabajo que hasta el momento han estado separadas: el razonamiento y el razonamiento estadístico, de tal manera que se gane en precisión y claridad conceptual. El primer paso para hacerlo es establecer una correlación entre el razonamiento y el razonamiento estadístico controlando los conocimientos previos. Dado que se sabe que los conocimientos previos influyen tanto en el razonamiento como el desempeño de otras áreas, en particular sobre la estadística, nos hemos propuesto despejar esta variable haciendo también su evaluación.

Pregunta de investigación

¿Existe correlación entre razonamiento condicional y razonamiento estadístico?

Hipótesis

A más alto el nivel de razonamiento más alto el puntaje en la prueba de razonamiento estadístico. A más conocimientos más alto el puntaje de la prueba de razonamiento estadístico.

La prueba decisiva será encontrar correlación entre altos niveles de razonamiento, bajos niveles de conocimientos y alto puntaje en la prueba de razonamiento estadístico. Esto indicaría precisamente que la capacidad de razonar es la que mejora el razonamiento estadístico. Al mismo tiempo, un bajo nivel de razonamiento con un alto nivel de conocimientos y un alto puntaje en razonamiento estadístico no avalaría la hipótesis.

Aquí lo importante es relacionar un alto nivel de razonamiento estadístico con las dos anteriores variables: razonamiento y conocimientos. Las siguientes son las combinaciones posibles de estas tres variables:

1. Todas están altas
2. Todas están bajas
3. El razonamiento estadístico está alto (niveles 3 y 4) y las otras dos están bajas
4. El razonamiento estadístico está alto y al menos una de las otras dos está alta y la otra baja (esta es la importante, siempre y cuando sea que la que está alta sea la prueba de razonamiento, en todas sus fases)
5. El razonamiento estadístico está bajo (niveles 1 y 2) y al menos una de las otras dos está alta y la otra baja
6. El razonamiento estadístico está bajo y las otras dos están altas

Método

Se llevó a cabo un estudio descriptivo correlacional de tipo cuantitativo, y transeccional porque las distintas evaluaciones se hicieron en un periodo específico de tiempo. Se realizó sobre una muestra a conveniencia de alumnos universitarios.

Participantes

Una muestra definitiva de 47 alumnos de licenciatura (38 mujeres y 9 hombres), inicialmente 27, y en una segunda etapa se incorporaron otros 25 alumnos con edades entre 18 y 26 años ($M= 19.9$, $DE= 1.35$), todos provenientes de una universidad pública. La selección fue de modo intencional, previo consentimiento y garantía de confidencialidad y anonimato. De la muestra inicial se eliminaron cinco participantes después de aplicar el criterio de inclusión, que consistió en responder correctamente dos preguntas básicas de comprensión de las instrucciones.

Instrumentos

Se aplicaron cinco pruebas en total: una de conocimiento en estadística, tres para evaluar el razonamiento condicional en su nivel concreto, intermedio y abstracto; y por último, una prueba para valorar el razonamiento estadístico.

El *conocimiento en estadística* se evaluó mediante la adaptación de solo ocho reactivos originales de la prueba BLIS (*Basic Literacy In Statistics*),³ Ziegler (2014); los reactivos plantean cuestiones relativas al tema de la muestra estadística, y las respuestas son de opción múltiple, ejemplo:

El Instituto Nacional de Estadística y Geografía (INEGI) realizó una encuesta sobre el uso de tecnologías de la información a un grupo representativo a nivel nacional de 1,002 adultos mexicanos en el año 2000. De estos adultos, 21% tenía una cuenta de correo electrónico. Ahora, identifique la población sobre la cual el INEGI puede hacer inferencias a partir de los resultados de la encuesta y aclare cuál es la muestra de esa población:

- La población son todos los adultos mexicanos en 2000. La muestra corresponde a los 1,002 adultos mexicanos encuestados.
- La población son todos los adultos mexicanos en 2000. La muestra es 21% de los adultos mexicanos quienes poseen una cuenta de correo electrónico.
- La población son los 1,002 adultos mayores de 18 años encuestados. La muestra es todos los adultos mexicanos en 2000.

El *razonamiento lógico condicional* se midió con tres pruebas; su elaboración partió de la prueba original de las cuatro tarjetas de Wason (1968), misma que se utilizó y la denominamos prueba *abstracta*; la segunda prueba, sería la versión *concreta* y una tercera prueba extra que elaboramos con la intención de ubicarla entre la versión *abstracta* y *concreta*, la llamamos prueba *intermedia*. La quinta prueba evaluó el *razonamiento estadístico* alrededor del concepto de muestra estadística, mediante preguntas abiertas con base en la observación de un video corto.

Procedimiento

Se realizaron dos sesiones experimentales. En la primera se hizo la presentación a 27 participantes y se aplicó la primera prueba, evaluación de los conocimientos en estadística; después de terminar se les dieron instrucciones y se presentó un video con dos procedimientos de distribución muestral (alrededor de 8 minutos de duración y protagonizado por los autores). En una bolsa negra y opaca se colocaron 200 fichas (100 rojas y 100 azules); en el primer procedimiento se sacaron dos fichas al azar, se registraba el resultado en una pizarra y luego se regresaban a la bolsa, este procedimiento se repitió en 50 ocasiones; el segundo procedimiento consistió en la misma mecánica, pero ahora se sacaron siete fichas al azar, en 50 ensayos; al final del video se presentó a los participantes los resultados de esta distribución muestral con reemplazo en dos gráficas, una para cada procedimiento. La evaluación de las respuestas abiertas se realizó apegada a los lineamientos de la taxonomía SOLO de Biggs y Collins (1982, 1989). Una vez que terminaron, se

³ Previo permiso de la autora Laura Ann Ziegler.

les entregaron las demás pruebas de razonamiento condicional (*concreto* y *abstracto*). En una segunda sesión se incluyeron los 25 participantes restantes y se llevó a cabo el mismo procedimiento que el grupo anterior, solo que en este grupo se incluyó una prueba de razonamiento condicional *intermedia*. Todas las pruebas estaban impresas en hojas blancas y respondidas a mano.

Resultados

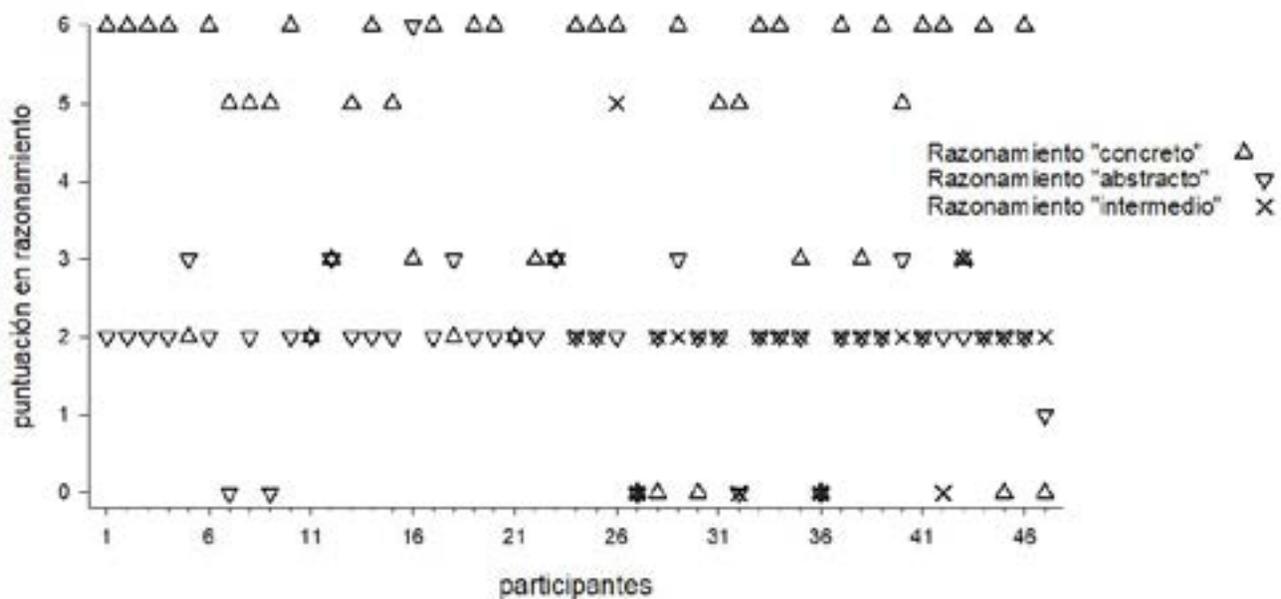
Los resultados de las correlaciones muestran una covariación positiva entre las pruebas de razonamiento abstracto y la prueba de razonamiento intermedio altamente significativa. Además, se encontró una relación negativa entre las puntuaciones de conocimiento en estadística y el puntaje en razonamiento abstracto.

Tabla 1. Matriz de correlaciones. Rho de Spearman

	Razonamiento estadístico	Conocimiento en estadística	Razonamiento <i>concreto</i>	Razonamiento <i>intermedio</i>	Razonamiento <i>abstracto</i>
Razonamiento estadístico	X				
Conocimiento en estadística	0.06	X			
Razonamiento <i>concreto</i>	-0.008	-0.20	X		
Razonamiento <i>intermedio</i>	0.30	0.06	0.21	X	
Razonamiento <i>abstracto</i>	0.08	-0.35*	0.05	0.54**	X

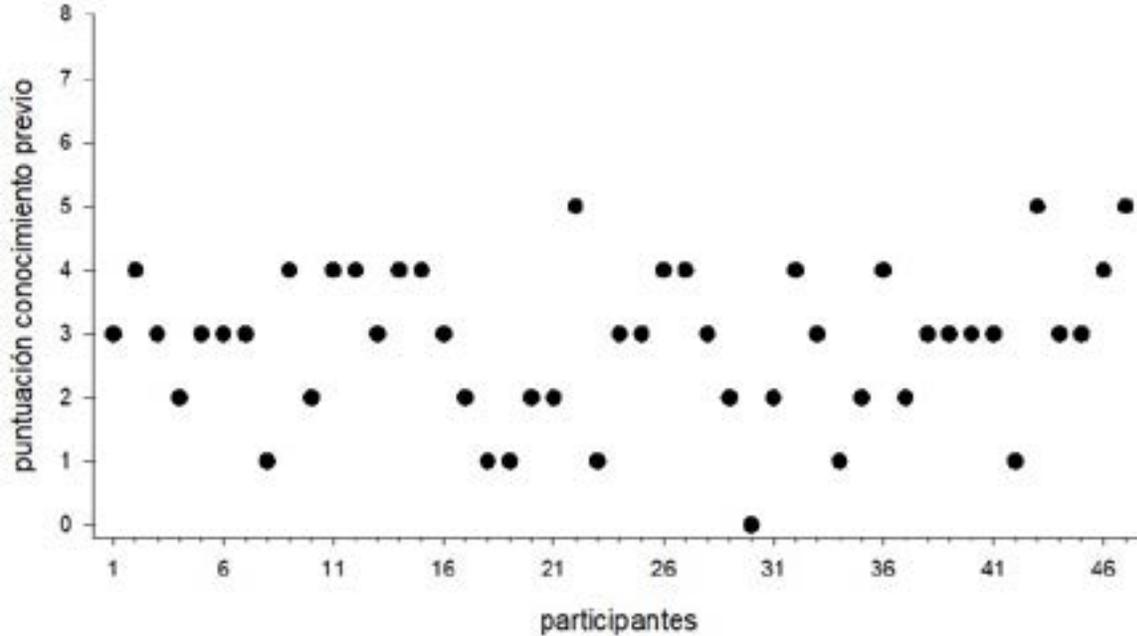
Se realizaron dos procedimientos de correlación parcial; el primero fue para evaluar la hipótesis nula de que no existe relación significativa entre los niveles de razonamiento estadístico y razonamiento concreto, después de controlar los efectos de conocimiento previo en estadística (N = 47).

Figura 1. Puntos para cada una de las pruebas de razonamiento condicional por participante



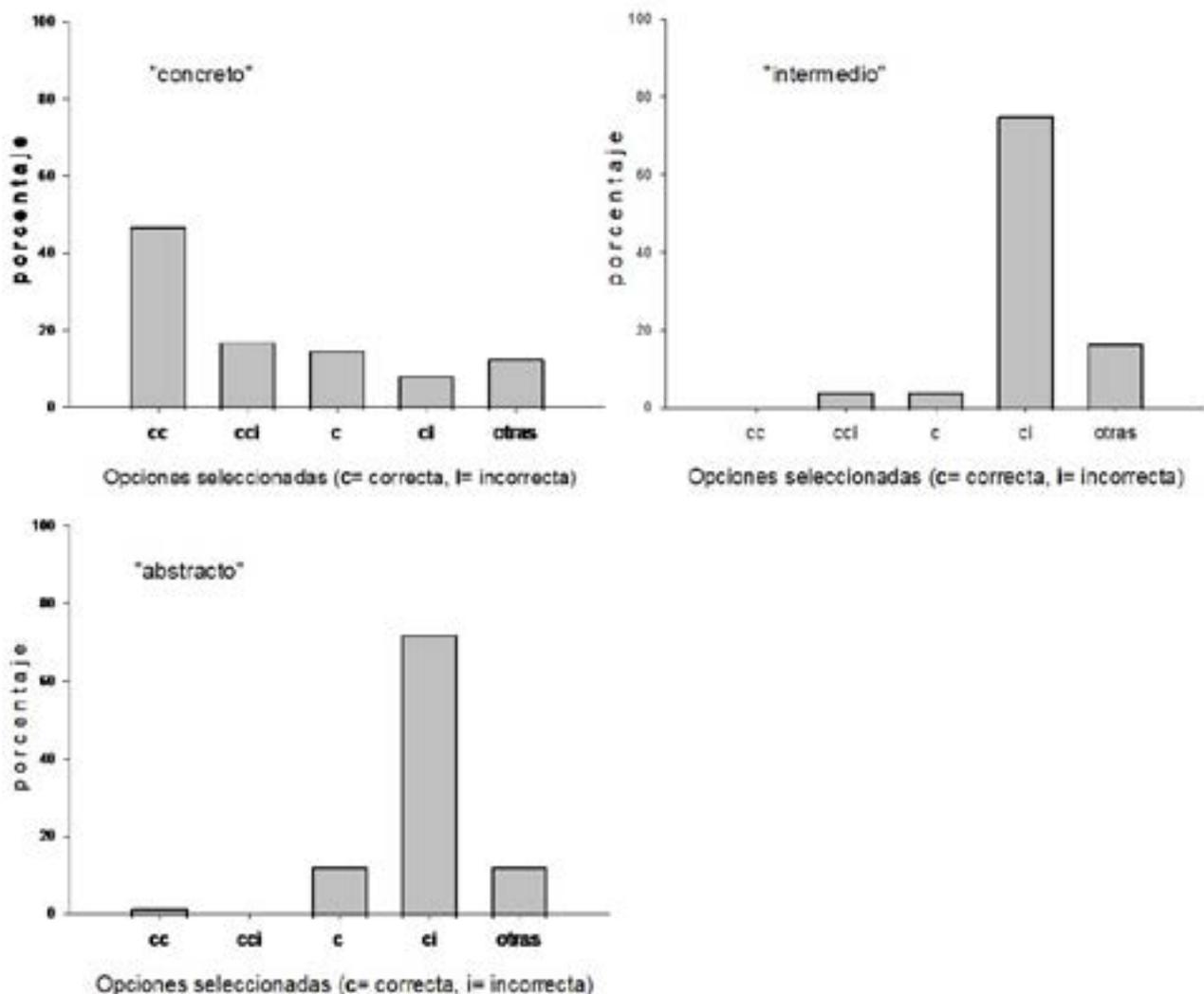
La figura 1 describe los puntos por participante en cada una de las pruebas de razonamiento condicional, se puede observar, para el razonamiento concreto, que la mayoría de los participantes obtuvo altas calificaciones, 5 y 6 puntos; para las pruebas de razonamiento abstracto e intermedio hay un rendimiento similar, 2 y 3 puntos, que es bajo.

Figura 2. Puntuación en conocimiento previo en estadística por participante



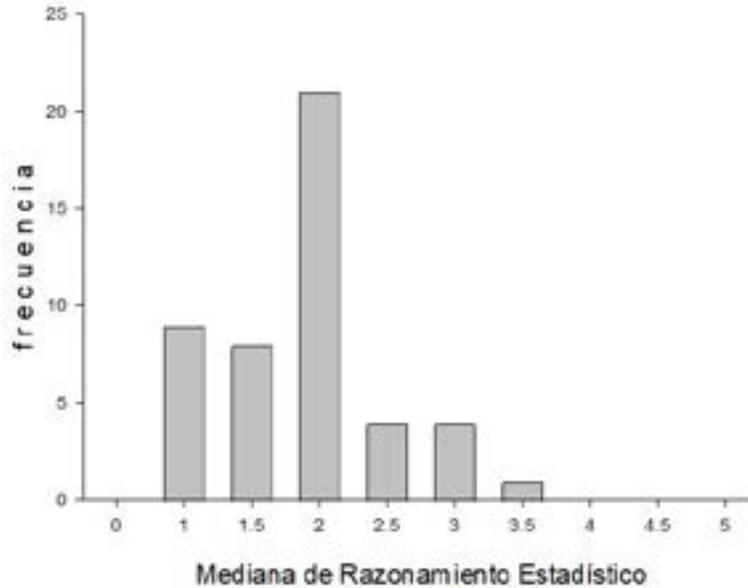
En la figura 2 se observa un pobre desempeño en conocimientos previos en estadística en todos los participantes, de manera particular, sobre temas relativos a la muestra estadística. Además, se evaluó el conocimiento estadístico de acuerdo con la taxonomía SOLO, con lo cual se confirma un bajo nivel de desempeño, situándose la mayoría de los participantes en un nivel unidimensional.

Figura 3. Razonamiento condicional (concreto, intermedio y abstracto). Muestra el porcentaje combinado de tarjetas seleccionadas



La ejecución en las tres pruebas de razonamiento condicional, en general, replican los resultados de Wason (1968). En la prueba concreta, la mitad de los participantes logran acertar las dos opciones correctas y en las pruebas intermedia y abstracta tienen resultados bastante similares, lo cual explica su correlación positiva en la tabla 1. En estas dos pruebas los participantes seleccionan, como lo encontró Wason, las tarjetas de confirmación bidireccional de la regla y no eligen la tarjeta de negación de la regla. En general, del total de participantes que seleccionaron la opción p , que implica confirmar la regla, solo 57% seleccionó como segunda opción q que también implica confirmar la regla de manera reversible, pero el problema es que la regla no es reversible: " p entonces q " y falso pensar que es igual " q entonces p ".

Figura 4. Frecuencia de la mediana de razonamiento estadístico, de acuerdo con la taxonomía SOLO



Se muestra la mediana del nivel de razonamiento estadístico que alcanzaron todos los participantes. Como puede observarse, el nivel es bajo, poco más de veinte participantes apenas logran un nivel 2 o unidireccional, y otro tanto similar en el nivel preestructural; pocos (cerca de 5) llegan a un nivel de razonamiento multidimensional SOLO.

Discusión y conclusiones

¿Es el razonamiento estadístico en realidad un razonamiento formal, tal y como lo definen las investigaciones en el área? La respuesta directa es no, si es que utilizamos como punto de referencia las investigaciones que se han realizado hasta el momento. Entonces, ¿cómo es que podemos decir que eso es razonamiento? En nuestra opinión, la mejor manera de comenzar a ligar las dos líneas de investigación es vincular los niveles de desempeño entre ambas.

Es en ese tenor que hemos aplicado tres pruebas de razonamiento en tres niveles de abstracción, siguiendo lo que otros investigadores en el área del razonamiento habían realizado antes. En primer lugar, aplicamos una versión concreta de la evaluación del razonamiento condicional a la que denominamos “el restaurante”, la cual formalmente tiene la misma estructura que la prueba abstracta, que es la versión de las cuatro tarjetas de Wason, la tercera de nuestras pruebas. A medio camino entre ambas, en lo que se refiere al nivel de abstracción, se aplicó una prueba a la que llamamos “las cartas de uno”.

Lo que nos encontramos en la prueba concreta es que una mayoría significativa de sujetos contestaron de manera correcta los cuestionamientos que les realizamos (¿Cuántas y a cuáles personas se les debe investigar para saber si pueden beber alcohol en un restaurante?); la mayoría identificó que solo se les debería investigar a dos: para saber si eran o no mayores de edad y si estaban tomando o no alcohol. Más allá de estas dos condiciones, no tiene sentido indagar más puesto que si sabes que alguien es mayor de edad no importa que esté o no tomando alcohol y, de la misma manera, no tiene sentido preguntar a las personas su edad si sabemos que están tomando una bebida sin alcohol.

Este dato está en consonancia con estudios que han demostrado que las versiones concretas de pruebas de razonamiento pueden ser contestadas de manera correcta. Se ha observado que se mejora la respuesta en la tarea de Wason si es presentada en términos menos abstractos (Johnson-Laird *et al.*, 1972), o cuando el material a ser juzgado tiene contenidos significativos (Griggs, Cox, 1982). Aunque pacientes con lesiones en la corteza prefrontal ventromedial no muestran esta mejora (Adolphs *et al.*, 1996).

Es decir, el contenido con que se presentan los problemas influye en la respuesta de los participantes. Así tendemos a centrarnos en la certeza o falsedad de las afirmaciones individuales del silogismo (Markovits, Nantel, 1989); razonamos basados en nuestras creencias (Klauer *et al.*, 2000) y entre más creíbles sean las afirmaciones con las que estamos razonando, más se ignora la estructura lógica (Evans *et al.*, 1983).

Datos muy diferentes son los que nos encontramos en la prueba de las cuatro tarjetas y las “cartas de uno”. Los datos de ambas pruebas casi son una calca una de la otra. En ambas, los datos nos arrojan una combinación de opciones que es claramente más elegida que el resto de las opciones, que es una combinación de una tarjeta correcta y una incorrecta, en ese orden, pero en la cual, además, no se eligió una tarjeta incorrecta cualquiera, sino aquella que tiende a confirmar y no a rechazar la hipótesis con la que se está trabajando en la solución de esos problemas.

Lo que demuestra la tarea de selección de Wason es la dificultad en las personas para razonar correctamente. Oaksford y Chater (1994) mostraron que 90% escoge E, que está bien; 60% escoge 4, que está mal; 25% elige 7, que está bien; y 15% D, que está mal. Solo 10% escoge E y 7. Es decir, las personas raramente fallan en aceptar el *modus ponens*, pero la frecuencia en que se acepta el *modus tollens* es solo un poco más grande que la frecuencia en que se aceptan inferencias inválidas (Evans, 1993). Demuestra también que hay un sesgo de confirmación de hipótesis y que raramente las personas tratan de falsear lo que piensan (esta es una de las razones por las que el falsacionismo de Karl Popper es importante).

Esto no quiere decir que las personas nunca razonan de manera correcta. Sí se puede razonar bien. De hecho, se ha demostrado que cuando las personas están haciendo una deducción correctamente se activa el lóbulo parietal (Qin *et al.*, 2003). Es solo que cuesta trabajo, tan así que aún el entrenamiento en lógica no mejora el desempeño en la tarea de Wason (Cheng *et al.*, 1986).

Pero el hecho de que los datos de ambas pruebas (las tarjetas de uno y las tarjetas de Wason) sean prácticamente idénticas, nos indica una de dos cosas o ambas a la vez. Una, que la prueba que nosotros diseñamos en realidad no es una prueba de razonamiento semiabstracto, tal y como era nuestro objetivo; dos, que a las personas se les deteriora su razonamiento casi inmediatamente que la situación se vuelve, aunque sea poco abstracta, fuera de su vida diaria. Esa es una cuestión que se deberá investigar en posibles investigaciones futuras.

Una cuestión más respecto a las pruebas de razonamiento es que los datos encontrados con la prueba de las cuatro tarjetas de Wason replican prácticamente igual otras investigaciones realizadas anteriormente (Oaksford, Chater, 1994).

A continuación, aplicamos la prueba de conocimientos en estadística. Los resultados en general nos muestran que los sujetos se encuentran en los niveles bajos ya que solo tres sujetos obtuvieron 5 de 8 puntos posibles; todos los demás estuvieron debajo de la mitad de los puntos a obtener (11 con 4 puntos, 17 con 3, 9 con 2, 6 con 1 y 1 con 0). Todos los sujetos, al ser alumnos de licenciatura, contaban con cursos previos de estadísticas aunque ya habían pasado años desde entonces.

Tomamos a estos sujetos en estas condiciones (ya con un curso de estadística, pero cursado hace tiempo) precisamente porque nos ayuda a comprobar nuestra hipótesis. Si hubiéramos elegido sujetos sin cursos de estadística se nos señalaría, con mucha razón, que esos no eran un grupo adecuado para estudiar el razonamiento estadístico si no tenían ni idea de que era de lo que les estábamos hablando. Pero tampoco era nuestro interés tener sujetos que acabaran de tomar un curso de estadística, porque la inmediatez del curso haría que estuviéramos midiendo más bien la efectividad del curso.

¿Por qué se desempeñaron mal los sujetos de estudio en nuestra prueba de conocimientos sobre estadística? Como decíamos anteriormente, su curso previo había sido hace mucho tiempo y su curva de aprendizaje indicaría que ya ha pasado el tiempo suficiente como para que se olvidara lo que habían aprendido. Otra razón es que las estrategias didácticas usadas no promovieron realmente el aprendizaje (Horton, 2013). También, probablemente, el curso previo no cambió los sesgos y concepciones previas y de sentido común con que llegan los estudiantes a los cursos (tales como la heurística de representatividad y la de disponibilidad, la ley de los pequeños números, la falacia del jugador, el sesgo de equiprobabilidad y la falacia de correlación, entre otras) (Garfiel, Ben-Zvi, 2007) y aunque aprendieron los conceptos y las habilidades que se les procuraron en su estudio previo, simplemente no los supieron utilizar de nuevo, es decir, no lograron la competencia estadística (Garfield, 1995).

La última serie de mediciones realizadas fueron sobre el razonamiento estadístico. Dado que esta era la parte central de la investigación, decidimos seguir la metodología aplicada por la mayoría de los investigadores en el área. Es decir, elegimos un problema estadístico (muestreo), definimos cuáles eran los elementos clave para resolver este problema (tamaño de la población,

nivel de confianza, margen de error, etc.), les hicimos una serie de preguntas relacionadas respecto al tamaño de la muestra (cuál muestra podía ser mejor para saber la proporción de fichas rojas o azules), analizamos sus respuestas para saber si estaban utilizando uno, dos o más o ningún elemento clave para resolver el problema, y con base en eso los clasificamos en la escala SOLO (nivel preestructural si su respuesta no utilizaba ni un solo elemento clave, uniestructural si usaba uno, multiestructural si manejaba dos o más, y relacional si integraba esos dos o más elementos clave).

Lo que encontramos es que los sujetos se encuentran en los niveles más bajos de la escala SOLO, aunque principalmente se encuentran en los niveles 1 y 2, preestructural y uniestructural respectivamente, muy pocos en el nivel 3, multiestructural y ninguno en el nivel 4, relacional.

Recordemos algunos de los estudios ya citados anteriormente (Jones *et al.*, 2000; Jones *et al.*, 2001; Watson *et al.*, 1995; Mooney, 2002; Li, Goos, 2013, Juárez, Inzunza, 2014): todos encontraron el mismo dato que nosotros: la mayoría de los sujetos fueron clasificados en los niveles bajos de la escala.

Dados estos pasos previos, entramos ya de lleno en la parte sustantiva del estudio: un análisis de correlación entre todas las mediciones. Recordemos que nuestras hipótesis son las siguientes: que tanto el razonamiento como el conocimiento tienen influencia en el razonamiento estadístico, pero dado que el razonamiento es nuestra variable de interés y que el conocimiento previo es una variable extraña que tratamos de controlar, la prueba decisiva sería encontrar una correlación entre altos niveles de razonamiento, junto con bajos niveles de conocimiento y un alto nivel de puntaje en la prueba de razonamiento estadístico. Esto indicaría que la capacidad de razonar es la que mejora el razonamiento estadístico. Al mismo tiempo, un bajo nivel de razonamiento con un alto nivel de conocimientos y un alto puntaje en razonamiento estadístico no avalaría la hipótesis, porque lo que estaría relacionado serían los conocimientos y no el razonamiento a un incremento en la ubicación de la escala SOLO.

En el análisis de correlación hay tres datos significativos: hay una correlación negativa entre razonamiento concreto y conocimiento previo, hay una correlación negativa entre conocimiento previo y razonamiento abstracto y hay una correlación negativa entre razonamiento concreto y razonamiento estadístico. Es decir, lo que nos dicen los dos primeros datos es que el conocimiento previo interfiere con el razonamiento tanto abstracto como concreto. Y el último dato nos dice que el razonamiento concreto está contrapuesto al razonamiento estadístico.

Veamos en primer lugar cómo el conocimiento previo interfiere con el razonamiento. Ya anteriormente habíamos abordado esta cuestión aseverando que los conocimientos previos ayudaban a resolver problemas de razonamiento, y que por ello las pruebas de razonamiento que apelan a cuestiones concretas se resuelven más fácilmente que las abstractas. Pero en esta ocasión estamos ante una situación completamente al revés.

A este tipo de errores se les conoce como errores de contenido; es decir es en aquellas ocasiones en las que nos centramos en la certeza o falsedad de las afirmaciones, por ejemplo, de un

silogismo (Markovits, Nantel, 1989) y por eso se da una preferencia basada en la creencia (Klauer *et al.*, 2000), es decir, entre más creíble sea una afirmación más se ignora la estructura lógica del problema que se está resolviendo. A los problemas familiares se les juzga por su solidez y no por su validez, por ello están influenciadas por la credibilidad del argumento (Evans *et al.*, 1983).

Este proceso tiene bases neurofisiológicas. Por ejemplo, se ha estudiado qué pasa en el momento de hacer comprobaciones científicas (Fungelsang, Dumber, 2005) y los resultados sugieren que el cerebro recluta mecanismos de aprendizaje cuando evalúa datos que son coherentes con hipótesis preexistentes y para reclutar mecanismos de detección de errores cuando evalúa datos que no son coherentes. Y también se ha visto que cuando las personas razonan en silogismos sin contenido se activan regiones parietales, pero si se juzgan materiales con contenido se activan regiones prefrontales y temporo-parietales (Goel *et al.*, 2000).

Hay varias propuestas que tratan de explicar esta situación. En primer lugar, está la que se conoce como sistemas duales de razonamiento, propuesta por Sloman (1996), en la que se proponen dos sistemas, uno asociativo que usa relaciones de temporalidad y similaridad, y un sistema basado en reglas que exhibe productividad, sistematicidad y composicionalidad, es decir, si se cumplen condiciones se aplica la regla y que, por lo tanto, tiene que ver con la deliberación, explicación, la adscripción de propósitos, etcétera.

Otra versión de sistemas duales de razonamiento es la de Tversky y Kahneman (1983), quienes proponen que en el razonamiento se utilizan más bien heurísticas y sesgos que les hacen caer en errores lógicos como la falacia de conjunción y que, a pesar de que se les explica a los sujetos en qué consiste el error y lo entienden, siguen repitiéndolo.

La crítica a estas propuestas de sistemas duales afirma que dos respuestas conflictivas no implican necesariamente dos sistemas de razonamiento (Gigerenzer, Rieger, 1996). Byrne (1989) hizo una investigación que ilustra que el razonamiento humano es sensible a cosas más allá de las reglas de inferencia y que las personas tienen un rico entendimiento del mundo real, que puede hacerlos caer en errores lógicos, pero que son las decisiones correctas en el mundo real. En el mismo tenor, Johnson-Laird (1983) afirma que las personas tienden a juzgar la verdad de una conclusión en el mundo real y no entienden qué es la validez de un argumento y simplemente juzgan si una conclusión es posible dadas las premisas.

Hasta aquí lo que tiene que ver con los dos primeros datos significativos encontrados en el análisis de correlación. El tercer dato significativo, que existe una correlación negativa entre el razonamiento concreto y el razonamiento estadístico, es probable que semejante situación tenga que ver con que el razonamiento estadístico implica necesariamente un razonamiento abstracto.

Como habíamos observado con anterioridad al discutir sobre el razonamiento estadístico y cómo es que se ha utilizado el modelo SOLO de Biggs y Collis, la mayoría de las investigaciones ubican a sus participantes en los niveles bajos de la taxonomía, es decir, el razonamiento estadístico en sus niveles más altos, no es fácil de alcanzar.

A esto debemos sumarle que el razonamiento estadístico es una parte del pensamiento estadístico, y la definición misma de lo que se entiende por pensamiento estadístico nos da la idea de lo difícil que es llegar a un alto grado de abstracción. Tomemos como ejemplo la definición propuesta por Ben-Zvi y Garfield (2004), quienes dicen que el pensamiento estadístico implica una comprensión de por qué y cómo las investigaciones que usan estadísticas se llevan a cabo y las “grandes ideas” que les subyacen. Estas ideas incluyen la naturaleza omnipresente de la variación y cuándo y cómo utilizar los métodos apropiados de análisis de datos, tales como resúmenes numéricos y visualizaciones de datos. El pensamiento estadístico implica una comprensión de la naturaleza del muestreo, cómo hacemos inferencias de las muestras a las poblaciones, y por qué se necesitan experimentos diseñados para establecer la causalidad. Incluye una comprensión de cómo se usan los modelos para simular fenómenos aleatorios, cómo se producen los datos para estimar probabilidades, y cómo, cuándo y por qué las herramientas inferenciales existentes pueden ser utilizadas para ayudar en el proceso de investigación. El pensamiento estadístico también incluye ser capaz de comprender y utilizar el contexto de un problema al momento de diseñar investigaciones y sacar conclusiones, y reconocer y comprender todo el proceso (desde la formulación de la pregunta hasta la recopilación de datos, la elección de análisis y la prueba suposiciones, etc.). Finalmente, los pensadores estadísticos deben ser capaces de criticar y evaluar los resultados de un problema resuelto o un estudio que usó estadísticas.

Entonces, nuestros sujetos como grupo ¿obtuvieron una calificación alta en razonamiento? Sí, pero solo cuando se les pidió que razonaran de manera concreta. Vamos a ponerlo como que es un punto a favor de nuestra hipótesis. Lo siguiente es, ¿tuvieron bajos niveles de conocimiento? Y la respuesta claramente es sí. Entonces aquí están puestas ya las condiciones causales que estábamos investigando (alto razonamiento y bajo conocimiento), y ahora solo falta ver si obtuvieron un alto nivel de razonamiento estadístico, y la respuesta es que no. Como grupo, nuestros sujetos obtuvieron un bajo nivel de razonamiento estadístico; de hecho, encontramos una correlación negativa entre razonamiento concreto y razonamiento estadístico. Básicamente, lo que nos dice es que a más alto el puntaje en el razonamiento concreto, menos puntos se van a obtener en la prueba de razonamiento estadístico.

Más allá de que los datos obtenidos con nuestra investigación no hayan sido del todo a favor de nuestras hipótesis, se hace necesario abrir un espacio de reflexión sobre nuestra investigación. En primer lugar, está la cuestión del desarrollo de las habilidades de razonamiento en entornos educativos. Existen varias propuestas sobre diferentes maneras de fortalecer el razonamiento estadístico, por ejemplo, el aprendizaje basado en el cerebro (Susilawati, Abdullah, 2020), la cual usa estrategias metacognitivas para facilitar procesos de conflicto, reflexión y descubrimiento. O bien, técnicas de cuestionamiento que permiten a los estudiantes reflexionar sobre sus ideas y conectar conceptos de manera significativa (Johnny *et al.*, 2015). El uso de diagramas y modelos para visualizar relaciones entre los elementos de un problema ha de-

mostrado también mejorar la capacidad de los estudiantes para analizar y resolver situaciones diversas (Yazmurotovna, 2024). Finalmente, involucrar a los estudiantes en actividades constructivas que promuevan la argumentación científica y la generación de hipótesis fortalece las habilidades de razonamiento (Engelmann *et al.*, 2016).

En segundo lugar, los sesgos cognitivos y las concepciones previas son factores determinantes que interfieren significativamente en el razonamiento estadístico, lo que tiene implicaciones directas para la educación en áreas que demandan pensamiento crítico y analítico. Diversos estudios han señalado cómo estos elementos distorsionan la comprensión de conceptos fundamentales y limitan la capacidad de los estudiantes para integrar nuevos conocimientos en sus esquemas mentales preexistentes. Por ejemplo, Busom i Piquer y López-Mayán (2015) demostraron que las concepciones erróneas y los sesgos cognitivos afectan la capacidad de los estudiantes para comprender y aplicar razonamientos en el ámbito económico. Friedman y Leslie (2004) identificaron cómo los niños muestran un sesgo sistemático al realizar predicciones basadas en creencias falsas, lo cual afecta su razonamiento. En el ámbito del pensamiento crítico, la influencia de los sesgos es igualmente evidente. Beaulac y Kenyon (2014) subrayaron que las concepciones previas moldean las percepciones de nueva información, lo que puede resultar en juicios defectuosos si no se corrigen a tiempo. Por último, Friedman (2017) aportó evidencia sobre cómo los sesgos afectan la percepción de aleatoriedad y el razonamiento probabilístico, distorsionando expectativas relacionadas con resultados inciertos. En conjunto, estos estudios subrayan la necesidad de enfoques educativos que promuevan la reflexión sobre los propios procesos de pensamiento. Esto implica no solo enseñar conceptos técnicos, sino también desafiar y dismantelar concepciones previas y sesgos que obstaculizan el desarrollo del razonamiento crítico y estadístico.

En tercer lugar, el razonamiento estadístico desempeña un papel fundamental en la vida cotidiana al facilitar la toma de decisiones informadas y mejorar las habilidades para resolver problemas en contextos diversos. Según Ningsih (2020), una comprensión sólida de las estadísticas permite a las personas enfrentar situaciones prácticas con mayor eficacia, ya que estas habilidades son esenciales para evaluar datos y tomar decisiones basadas en evidencia. Por su parte, Rahmatina *et al.* (2022) destacan que el razonamiento estadístico conecta conceptos e ideas de manera significativa, ayudando a los estudiantes a integrar diferentes fuentes de información y utilizarlas para tomar decisiones cotidianas. El uso de datos de la vida real en la enseñanza de la estadística mejora significativamente las habilidades de razonamiento al contextualizar los conceptos y hacerlos relevantes para los problemas que enfrentan las personas en su vida diaria. Así pues, preparar a los estudiantes para que comprendan y apliquen conceptos estadísticos en contextos prácticos es una meta central de la educación, ya que estos conocimientos inciden directamente en su capacidad para tomar decisiones informadas, enfrentar problemas y adaptarse a las demandas de un entorno complejo y dinámico.

Referencias

- Adolphs, R.; D. Tranel; A. Bechara; H. Damasio; A. Damasio (1996). Neuropsychological Approaches to Reasoning and Decision-Making. In *Neurobiology of Decision-Making*. Germany: Springer-Berlin-Heidelberg, 157-179.
- Al-Labadi, L.; Z. Baskurt; M. Evans (2018). Statistical Reasoning: Choosing and Checking the Ingredients, Inferences Based on a Measure of Statistical Evidence with some Applications. *Entropy*, 20(4), 289. <https://www.mdpi.com/1099-4300/20/4/289>
- Ausubel, D.; J. Novak; H. Hanesian (1976). Significado y aprendizaje significativo. *Psicología educativa: Un punto de vista cognoscitivo*, 1(2), 53-106. <https://cmapspublic2.ihmc.us/rid=1J3D72HXW-2CBD02Q-PW5/aprendizaje%20Ausubel.pdf>
- Azmay, N.; R. Rosli; S. Maat; M. Mahmud (2023). Educational Research Trends on Statistical Reasoning and Statistical Thinking: A Systematic Literature Review. *International Journal of Academic Research in Progressive Education and Development*, 12(2), 586-600. <https://ijarped.com/index.php/journal/article/view/1198>
- Beaulac, G.; T. Kenyon (2014). Critical Thinking Education and Debiasing (AILACT Essay Prize Winner 2013). *Informal Logic*, 34(4), 341-363. <https://doi.org/10.22329/il.v34i4.4203>
- Ben-Zvi, D.; J. Garfield, J. (2004). *Statistical Literacy, Reasoning and Thinking: Goals, Definitions and Challenges*. En Ben-Zvi, D.; J. Garfield (eds.). *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*. Kluwer Academic Publishers, 3-15. https://link.springer.com/chapter/10.1007/1-4020-2278-6_1
- Biggs, J.; K. Collis (1982). *Evaluating the Quality of Learning: The SOLO Taxonomy*. USA: Academic Press.
- Biggs, J.; K. Collis (1989). Towards a Model of School-Based Curriculum Development and Assessment Using the SOLO Taxonomy. *Australian Journal of Education*, 33(2). 151-163. <https://doi.org/10.1177/168781408903300205>
- Burrill, G. (2020). Statistical Literacy and Quantitative Reasoning: Rethinking the Curriculum. *New Skills in the Changing World of Statistics Education: Proceedings of the Roundtable Conference of the International Association for Statistical Education (IASE), Held online*. <http://iase-web.org/documents/papers/rt2020/IASE2020 Roundtable 19 BURRILL.pdf?1610923749>
- Busom, I.; C. López-Mayán (2015). Student Preconceptions and Learning Economic Reasoning. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2704371
- Byrne, R. (1989). Suppressing Valid Inferences with Conditionals. *Cognition*, 31(1), 61-83. https://www.researchgate.net/publication/20449155_Suppressing_Valid_Inferences_With_Conditionals
- Chan, S.; Z. Ismail (2014). Developing Statistical Reasoning Assessment Instrument for High School Students in Descriptive Statistics. *Procedia-Social and Behavioral Sciences*, (116), 4338-4343. <https://doi.org/10.1016/j.sbspro.2014.01.943>

- Chan, S.; Z. Ismail; B. Sumintono (2015). The Impact of Statistical Reasoning Learning Environment: A Rasch Analysis. *Advanced Science Letters*, 21(5), 1211-1215. <https://www.researchgate.net/publication/283470216> The Impact of Statistical Reasoning Learning Environment A Rasch Analysis
- Chan, S.; Z. Ismail; B. Sumintono (2016). A Framework for Assessing High School Students' Statistical Reasoning. *PloS one*, 11(11), e0163846. <https://doi.org/10.1371/journal.pone.0163846>
- Chance, B.; J. Garfield; R. delMas (2000). Developing Simulation Activities to Improve Students' Statistical Reasoning. *Journal of Statistics Education*, 7(3). <https://www.researchgate.net/publication/283470216> The Impact of Statistical Reasoning Learning Environment A Rasch Analysis
- Cheng, P.; K. Holyoak; R. Nisbett; L. Oliver (1986). Pragmatic versus Syntactic Approaches to Training Deductive Reasoning. *Cognitive Psychology*, 18(3), 293-328. [https://doi.org/10.1016/0010-0285\(86\)90002-2](https://doi.org/10.1016/0010-0285(86)90002-2)
- delMas, R. (2002). Statistical Literacy, Reasoning, and Thinking: A Commentary. *Journal of Statistics Education*, 10(2). <http://dx.doi.org/10.1080/10691898.2002.11910679>
- delMas, R. (2004). A Comparison of Mathematical and Statistical Reasoning. In *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*. Dordrecht: Springer-Kluwer Academic Publishers, 79-95. https://link.springer.com/chapter/10.1007/1-4020-2278-6_4
- Rahmatina, D.; T. Nusantara; N. Parta; H. Susanto (2022). Statistical Reasoning Process of Students in Decision Making Using Commognitive Framework. *Acta Scientiae*, 24(3), 63-88. <https://doi.org/10.17648/acta.scientiae.6603>
- Duarte, J.; S. Cazares (2014). Comprensión y razonamiento de profesores de Matemáticas de bachillerato sobre conceptos estadísticos básicos. *Perfiles Educativos*, 36(146), 14-29. <https://doi.org/10.22201/iisue.24486167e.2014.146.46024>
- Engelmann, K.; B. Neuhaus; F. Fischer (2016). Fostering Scientific Reasoning in Education—Meta-Analytic Evidence from Intervention Studies. *Educational Research and Evaluation*, 22(5-6), 333-349. <https://doi.org/10.1080/13803611.2016.1240089>
- Evans, J.; J. Barston; P. Pollard (1983). On the Conflict between Logic and Belief in Syllogistic Reasoning. *Memory & Cognition*, 11(3), 295-306. <https://doi.org/10.3758/bf03196976>
- Evans, J. (1993). The Mental Model Theory of Conditional Reasoning: Critical Appraisal and Revision. *Cognition*, 48(1), 1-20. [https://doi.org/10.1016/0010-0277\(93\)90056-2](https://doi.org/10.1016/0010-0277(93)90056-2)
- Evans, M. (2019). The Measurement of Statistical Evidence as the Basis for Statistical Reasoning. *Proceedings*, 46(1), 7. <https://doi.org/10.3390/ecea-5-06682>
- Friedman, H. (2017). Cognitive Biases that Interfere with Critical Thinking and Scientific Reasoning: A Course Module. *Social Science Research Network*, 10. https://www.academia.edu/32667965/Cognitive_Biases_that_Interfere_with_Critical_Thinking_and_Scientific_Reasoning_A_Course_Module

- Friedman, O.; A. Leslie (2004). Mechanisms of Belief-Desire Reasoning: Inhibition and Bias. *Psychological Science*, 15(8), 547-552. <https://doi.org/10.1111/j.0956-7976.2004.00717.x>
- Fugelsang, J.; K. Dunbar (2005). Brain-Based Mechanisms Underlying Complex Causal Thinking. *Neuropsychologia*, 43(8), 1204-1213. <https://doi.org/10.1016/j.neuropsychologia.2004.10.012>
- Gal, I.; L. Ginsburg; C. Schau (1997). Monitoring Attitudes and Beliefs in Statistics Education. *The Assessment Challenge in Statistics Education*, 12, 37-51. https://www.researchgate.net/profile/Iddo-Gal/publication/333402737_Gal_Iddo_Garfield_Joan_1997_book_all_19_chapters_The_assessment_challenge_in_statistics_education/links/5cebeb4c299bf109da7330f4/Gal-Iddo-Garfield-Joan-1997-book-all-19-chapters-The-assessment-challenge-in-statistics-education.pdf
- Galotti, K. (1989). Approaches to Studying Formal and Everyday Reasoning. *Psychological Bulletin*, 105(3), 331-351. <https://psycnet.apa.org/doi/10.1037/0033-2909.105.3.331>
- Garfield, J. (1995). How Students Learn Statistics. *International Statistical Review/Revue Internationale de Statistique*, 63(1), 25-34. <https://iase-web.org/documents/intstatreview/95.Garfield.pdf>
- Garfield, J.; I. Gal (1999). Teaching and Assessing Statistical Reasoning. *Developing Mathematical Reasoning in Grades K12*, 207-219. https://www.researchgate.net/publication/247700710_Teaching_and_assessing_statistical_reasoning
- Garfield, J.; B. Chance (2000). Assessment in Statistics Education: Issues and Challenges. *Mathematical Thinking and Learning*, 2(1-2), 99-125. https://doi.org/10.1207/S15327833MTL0202_5
- Garfield, J. (2003). Assessing Statistical Reasoning. *Statistics Education Research Journal*, 2(1), 22-38. <https://iase-web.org/ojs/SERJ/article/view/557/420>
- Garfield, J.; D. Ben-Zvi (2007). How Students Learn Statistics Revisited: A Current Review of Research on Teaching and Learning Statistics. *International Statistical Review*, 75(3), 372-396. <https://doi.org/10.1111/j.1751-5823.2007.00029.x>
- Gigerenzer, G.; T. Reiger (1996). How do We Tell an Association from a Rule? A Response to Sloman. *Psychological Bulletin*, 119(1), 23-26. https://pure.mpg.de/rest/items/item_2547830_4/component/file_2561651/content
- Goel, V.; C. Buchel; C. Frith; R. Dolan (2000). Dissociation of Mechanisms Underlying Syllogistic Reasoning. *Neuroimage*, 12(5), 504-514. <https://doi.org/10.1006/nimg.2000.0636>
- Gómez-Blancarte, A.; R. Chávez (2022). The Intersection of Statistical Literacy, Reasoning, and Thinking. *11th International Conference on Teaching Statistics (ICOTS11). Bridging the Gap: Empowering & Educating Today's Learners in Statistics*. http://iase-web.org/icots/11/proceedings/pdfs/ICOTS11_262_CHVEZAGU.pdf?1669865545
- Griggs, R.; J. Cox (1982). The Elusive Thematic - Materials Effect in Wason's Selection Task. *British Journal of Psychology*, 73(3), 407-420. <https://doi.org/10.1111/j.2044-8295.1982.tb01823.x>
- Hadar, N.; L. Henkin (1978). Children's Conditional Reasoning Part II: Towards a Reliable Test of Conditional Reasoning Ability. *Educational Studies in Mathematics*, 97-114. <https://doi.org/10.1007/BF00352195>

- Horton, N. (2013). I Hear, I Forget. I Do, I Understand: A Modified Moore-Method Mathematical Statistics Course. *The American Statistician*, 67(4), 219-228. <https://doi.org/10.1080/00031305.2013.849207>
- Husureau, T.; D. Cousineau; S. Zang; R. Gibeau (2024). A Compendium of Common Heuristics, Misconceptions, and Biased Reasoning Used in Statistical Thinking. *The Quantitative Methods for Psychology*, 20(1), 57-75. <https://www.tqmp.org/RegularArticles/vol20-1/p057/p057.pdf>
- Jansson, L. (1978). A Comparison of Two Approaches to the Assessment of Conditional Reasoning Abilities. *Journal for Research in Mathematics Education*, 9(3), 175-188. <https://doi.org/10.2307/748996>
- Johnny, J.; M. Abu; A. Abdullah; N. Atan; M. Mokhtar (2015). Reasoning Skills among Students: A Meta-analysis. *2nd International Education Postgraduate Seminar 2015 (IEPS 2015)* https://www.researchgate.net/publication/318205303_Reasoning_Skills_among_Students_a_Meta-Analysis
- Johnson-Laird, P.; P. Legrenzi; M. Sonio (1972). Reasoning and a Sense of Reality. *British Journal of Psychology*, 63(3), 395-400. <https://bpspsychub.onlinelibrary.wiley.com/doi/epdf/10.1111/j.2044-8295.1972.tb01287.x>
- Johnson-Laird, P. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Harvard University Press, 61(4). <https://doi.org/10.2307/414498>
- Jones, G.; C. Thornton; C. Langrall; E. Mooney; B. Perry; I. Putt (2000). A Framework for Characterizing Children's Statistical Thinking. *Mathematical Thinking and Learning*, 2(4), 269-307. https://doi.org/10.1207/S15327833MTL0204_3
- Jones, G.; C. Langrall; C. Thornton; E. Mooney; A. Wares; M. Jones; S. Nisbet (2001). Using Students' Statistical Thinking to Inform Instruction. *The Journal of Mathematical Behavior*, 20(1), 109-144. [https://doi.org/10.1016/S0732-3123\(01\)00064-5](https://doi.org/10.1016/S0732-3123(01)00064-5)
- Juárez, J.; S. Inzunza (2014). Comprensión y razonamiento de profesores de Matemáticas de bachillerato sobre conceptos estadísticos básicos. *Perfiles Educativos*, 36(146), 14-29. <https://doi.org/10.22201/iissue.24486167e.2014.146.46024>
- Kinnear, V. (2013). *Young Children's Statistical Reasoning: A Tale of Two Contexts*. Doctoral Dissertation. AUS: Queensland University of Technology. https://eprints.qut.edu.au/63496/1/Virginia_KINNEAR_Thesis.pdf
- Klauer, K.; J. Musch, J; B. Naumer (2000). On Belief Bias in Syllogistic Reasoning. *Psychological Review*, 107(4), 852. <https://psycnet.apa.org/doi/10.1037/0033-295X.107.4.852>
- Kuhn, D. (1977). Conditional Reasoning in Children. *Developmental Psychology*, 13(4), 342. <https://psycnet.apa.org/doi/10.1037/0012-1649.13.4.342>
- Li, K.; M. Goos (2013). Evaluating Students' Correlation Graphing Capability Using SOLO Taxonomy. *59th ISI World Statistics Conference (WSC)*. https://www.researchgate.net/publication/305984250_Evaluating_Students'_Correlation_Graphing_Capability_Using_SOLO_Taxonomy

- Lovett, M. (2001). A Collaborative Convergence on Studying Reasoning Processes: A Case Study in Statistics. *Cognition and Instruction: Twenty-five years of progress*, 347-384.
- Markovits, H.; G. Nantel (1989). The Belief-Bias Effect in the Production and Evaluation of Logical Conclusions. *Memory & Cognition*, 17(1), 11-17. <https://link.springer.com/article/10.3758/BF03199552>
- Márquez, R. (2012). *El desarrollo del pensamiento matemático y la actividad docente*. México: Gedisa.
- Mooney, E. (2002). A Framework for Characterizing Middle School Students' Statistical Thinking. *Mathematical Thinking and Learning*, 4(1), 23-63. http://dx.doi.org/10.1207/S15327833MTL0204_3
- Nicholson, J.; J. Ridgway, J.; S. McCusker (2006). Reasoning with Data—Time for a Rethink? *Teaching Statistics*, 28(1), 2-9. <http://dx.doi.org/10.1111/j.1467-9639.2006.0228a.x>
- Ningsih, Y.; R. Rohana (2020). Students' Statistical Reasoning in Statistics Method Course. *Online Submission*, 14(1), 81-90. https://www.researchgate.net/publication/338993534_STUDENTS'_STATISTICAL_REASONING_IN_STATISTICS_METHOD_COURSE
- Nolan, C.; E. Brandon (1984). *Conditional Reasoning in Jamaica*. Conference on Thinking, Harvard, Boston Mass. <https://uwispace.sta.uwi.edu/server/api/core/bitstreams/e1b4b819-98a2-4db7-becf-9dbbe7f000d9/content>
- Nor, N.; N. Idris (2010). Assessing Students' Informal Inferential Reasoning Using SOLO Taxonomy Based Framework. *Procedia-Social and Behavioral Sciences*, 2(2), 4805-4809. https://www.researchgate.net/publication/248607054_Assessing_Students'_Informal_Inferential_Reasoning_using_SOLO_Taxonomy_based_Framework
- Oaksford, M.; N. Chater (1994). A Rational Analysis of the Selection Task as Optimal Data Selection. *Psychological Review*, 101(4), 608-631. <https://pages.ucsd.edu/~cmckenzie/Oaksford-ChaterPsychReview1994.pdf>
- Qin, Y.; M. Sohn; J. Anderson; V. Stenger; K. Fissell; A. Goode; C. Carter (2003). Predicting the Practice Effects on the Blood Oxygenation Level-Dependent (BOLD) Function of fMRI in a Symbolic Manipulation Task. *Proceedings of the National Academy of Sciences*, 100(8), 4951-4956. <https://pmc.ncbi.nlm.nih.gov/articles/PMC153661/pdf/pq0803004951.pdf>
- Rahmatina, D.; T. Nusantara; N. Parta; H. Susanto (2022). Statistical Reasoning Process of Students in Decision Making Using Commognitive Framework. *Acta Scientiae*, 24(3), 63-88. <https://doi.org/10.17648/acta.scientiae.6603>
- Reading, C. (2002). Profile for Statistical Understanding. *Proceedings of the Sixth International Conference on Teaching Statistics, Cape Town, South Africa*. <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=1e88632259a4988a581bf8abe3d7a58185a716b2>
- Roberts, D.; J. Saxe (1982). Validity of a Statistics Attitude Survey: A follow-up Study. *Educational and Psychological Measurement*, 42(3), 907-912. <https://doi.org/10.1177/001316448204200326>
- SEP (2017). *Aprendizajes claves para la educación integral*. México: SEP. https://www.sep.gob.mx/work/models/sep1/Resource/10933/1/images/Aprendizajes_clave_para_la_educacion_integral.pdf

- Sloman, S. (1996). The Empirical Case for Two Systems of Reasoning. *Psychological Bulletin*, 119(1), 3-22. <https://psycnet.apa.org/doi/10.1037/0033-2909.119.1.3>
- Stacey, K. (2006). What Is Mathematical Thinking and why Is It Important. https://www.criced.tsukuba.ac.jp/math/apec/apec2007/paper_pdf/Kaye%20Stacey.pdf
- Stein, E. (1996). *Without Good Reason*. England: Clarendon Press. <https://doi.org/10.1017/S0031819100057193>
- Susilawati, W.; R. Abdullah (2020). Statistical Reasoning through Metacognitive Brain-Based Learning. *Journal Analisa*, 6(1), 40-46. <https://journal.uinsgd.ac.id/index.php/analisa/article/view/8158>
- Sutarso, T. (1992). Some Variables in Relation to Students' Anxiety in Learning Statistics. <https://files.eric.ed.gov/fulltext/ED353334.pdf>
- Tempelaar, D. (2004). Statistical Reasoning Assessment: An Analysis of the SRA Instrument. <https://www.rossmanchance.com/artist/proceedings/tempelaar.pdf>
- Tversky, A.; D. Kahneman (1983). Extensional versus Intentional Reasoning: The conjunction Fallacy in Probability Judgment. *Psychological Review*, 90(4), 293-315. <https://psycnet.apa.org/doi/10.1037/0033-295X.90.4.293>
- Vosniadou, S.; X. Vamvakoussi; I. Skopeliti (2008). *International Handbook of Research on Conceptual Change*. USA: Routledge.
- Wang, W.; X. Wang; G. Chen (2009). Survey and Analysis of the Statistical Reasoning among High School Students in China and Dutch. *Journal of Mathematics Education*, 2(1), 15-26. <https://doi.org/10.52041/serj.v2i1i.67>
- Wason, P. (1968). Reasoning about a Rule. *Quarterly Journal of Experimental Psychology*, 20(3), 273-281. <https://doi.org/10.1080/14640746808400161>
- Watson, J.; K. Collis; R. Callingham; J. Moritz (1995). A Model for Assessing Higher Order Thinking in Statistics. *Educational Research and Evaluation*, 1(3), 247-275. https://www.researchgate.net/publication/248914097_A_Model_for_Assessing_Higher_Order_Thinking_in_Statistics
- Watson, J.; J. Moritz (1998). The Beginning of Statistical Inference: Comparing Two Data Sets. *Educational Studies in Mathematics*, 37(2), 145-168. <https://doi.org/10.1023/A:1003594832397>
- Yazmurotovna, K. (2024). Improving the Logical Thinking System of Primary School Pupils. *International Journal of Artificial Intelligence*, 4(03), 673-674. <https://www.academicpublishers.org/journals/index.php/ijai/article/view/764>
- Ziegler, L. (2014). Reconceptualizing Statistical Literacy: Developing an Assessment for the Modern Introductory Statistics Course. University Digital Conservancy, <https://hdl.handle.net/11299/165153>