Niveles de razonamiento de estudiantes de secundaria y bachillerato en una situación-problema de probabilidad

Jaime I. García, Cinvestav – IPN (México)

Miguel Medina, Cinvestav – IPN (México)

Ernesto Sánchez, Cinvestav – IPN (México)

Recibido el 9 de mayo de 2014; aceptado el 8 de octubre de 2014

Niveles de razonamiento de estudiantes de secundaria y Bachillerato en una situación-problema de probabilidad.

Resumen

En el presente trabajo se explora la manera en que estudiantes de secundaria (14-15 años) y Bachillerato (15-16 años), que no han estudiado probabilidad, resuelven una tarea que involucra algunas nociones básicas de la teoría de probabilidades. El propósito es identificar los rasgos característicos de los niveles de razonamiento que exhiben los estudiantes cuando transitan de la secundaria al Bachillerato frente a una situación-problema de probabilidad. Se utiliza la taxonomía SOLO para organizar las respuestas en niveles estructurales. En general, los estudiantes de Bachillerato muestran un nivel más alto de razonamiento que los estudiantes de secundaria.

Palabras clave. Niveles de razonamiento; Taxonomía SOLO, Probabilidad; Distribución binomial

Níveis de razonamiento e abstracção de alunos do ensino básico e do ensino secundário numa situação-problema de probabilidades.

Resumo

O presente artigo explora a maneira como alunos do 3º ciclo do ensino básico (14-15 anos) e do ensino secundário (15-16 anos), que não estudaram probabilidades, resolvem uma tarefa que envolve algumas noções básicas da teoria de probabilidades. O propósito é identificar os traços característicos dos níveis de razonamiento que exibem os estudantes quando transitam da secundária ao Bachillerato em frente a uma situação-problema de probabilidade. Utiliza-se a taxonomia SOLO para organizar as respostas em níveis estruturais. Em geral, os alunos do ensino secundário mostram um maior nível de raciocínio do que os alunos do 3ºciclo do ensino básico, e as suas respostas distinguem-se pelo tipo de abstração que pressupõem.

Palavras chave. Níveis de razonamiento, SOLO, Probabilidade, distribuição binomial

Para citar: García, J. I., Medina, M., & Sánchez, E. (2014). Niveles de razonamiento de estudiantes de secundaria y Bachillerato en una situación-problema de probabilidad. *Avances de Investigación en Educación Matemática, 6*, 5- 23.

© Sociedad Española de Investigación en Educción Matemática (SEIEM). www.seiem.es

Levels of reasoning and abstraction of middle and high school students in a problem-situation of probability

Abstract

This paper explores the way in which middle school and high school students, who have not studied probability, solve a task that involves some basic notions of probability. The aim is to identify the characteristic features of the reasoning levels that middle and high school students exhibit when they face a probability situation-problem. The SOLO taxonomy is used to organize the responses in structural levels. In general, high school students showed a higher level of reasoning than middle school students.

Key words. Reasoning levels, SOLO taxonomy, Probability, Binomial distribution.

.**Niveaux de raisonnement et abstraction des étudiants du collège et du lycée dans une situation-problème de probabilité**

Résumé

Dans ce travail on explore la manière dont les élèves du collège et du lycée, qui ne jamais ont fait des cours de probabilité, abordent une tâche qui implique l'usage de certaines notions de base de la théorie des probabilités. L'objectif a été l'identification des traits caractéristiques des niveaux de raisonnement manifestés par les étudiants lorsqu'ils traversent de l'enseignement du collège au du lycée face à une situation-problème de probabilité. On a utilisé la taxonomie SOLO pour organiser les réponses selon les niveaux structurels. En général les lycéens montrent un niveau plus élevé de raisonnement que les élèves du collège.

Paroles clés. Niveaux de raisonnement, SOLO, Probabilité, Distribution binomiale

1. Introducción

La probabilidad es una disciplina cuyo aprendizaje puede ayudar a desarrollar una forma de pensamiento diferente de la que proporciona el aprendizaje de otras áreas de la matemática (Fischbein & Snarch, 1997). En efecto, la probabilidad emergió de la preocupación de resolver problemas asociados a situaciones de azar e incertidumbre (Hacking, 1975; Batanero, Henry & Parzysz, 2005), conceptos que son ajenos a los objetos de las demás ramas de la matemática. La probabilidad se ha ganado la reputación de ser una materia difícil y de generar numerosas y aparentes paradojas en estudiantes de todos los niveles de escolaridad (Borovcnik, 2012; Borovcnik & Bentz, 1991; Shaughnessy, 1992). Aunque la discusión acerca de cuál es el momento oportuno para iniciar la enseñanza de la probabilidad podría ser controvertida, la mayoría de países occidentales la han incluido en los currículos de matemáticas desde los niveles básicos, y su presencia continúa hasta llegar al nivel universitario. En consecuencia, los estudios didácticos sobre el desarrollo del pensamiento probabilístico pueden ser ubicados en cualquier nivel escolar, sólo variando en el tipo y nivel de complejidad de los problemas que se estudian. Al final de la secundaria (14-15 años) se suele considerar que los estudiantes son capaces de entender la relación entre las frecuencias relativas y la probabilidad; mientras que en Bachillerato (15-18 años) pueden entender los conceptos de variable aleatoria y distribución; lo cual se refleja en los currículos de varios países (Jones, Langrall & Mooney, 2007).

Los postulados constructivistas sobre el aprendizaje han sido ampliamente aceptados por la comunidad de investigadores educativos, en particular, por los educadores matemáticos y estadísticos (Garfield, 1995). La teoría constructivista sostiene que el aprendizaje es un proceso en el cual los sujetos construyen activamente su propio conocimiento (Von Glasersfeld, 1990). Así, el conocimiento que posee una persona no es algo elaborado externamente a él, que le haya sido transmitido mediante explicaciones, ni tampoco algo que el sujeto tenga de manera innata, sino que emerge de una combinación de la estructura del conocimiento que tiene el sujeto en conjunción con la actividad que realiza en ciertas condiciones favorables.

Con relación a la enseñanza, la teoría asume que su objetivo es proveer de oportunidades a los estudiantes para que realicen dicha construcción, y para hacerlo es necesario que el profesor conozca las dificultades que aquellos poseen y los niveles de razonamiento que son capaces de alcanzar. En este sentido, es útil la realización de estudios de diagnóstico sobre el pensamiento de los estudiantes en los temas en los cuales van a recibir instrucción; estudios que exploren y ofrezcan información que permita formular hipótesis acerca de los conocimientos que poseen y las dificultades que enfrentan. El objetivo del presente estudio es identificar algunos rasgos característicos de los niveles de razonamiento que exhiben los estudiantes cuando transitan de la secundaria al Bachillerato frente a una situación-problema de probabilidad. Para llevarlo a cabo, se exploran y analizan las respuestas de estudiantes de dos poblaciones: unos se encuentran estudiando al final de la secundaria y otros al principio del Bachillerato. Las respuestas se refieren a un cuestionario diagnóstico que se basa en una situación de probabilidad, cuya solución se relaciona con el concepto de distribución binomial.

2. Antecedentes

Los estudios directamente vinculados a la presente investigación se refieren a la distribución binomial; este concepto, sin embargo, se ubica en un sistema o red de conceptos que por sí mismos son objeto de estudios didácticos. De estos, aquí se hará un breve comentario de dos de ellos, a saber, aleatoriedad y espacio muestral; además se mencionan los estudios que se han encontrado en la literatura sobre distribución binomial.

El primero de esos conceptos es el de *aleatoriedad*, del cual se encuentra un análisis desde una perspectiva didáctica en Batanero, Serrano y Green (1998). Estos autores afirman que es fundamental en el estudio de la probabilidad, ya que se requiere desarrollar una comprensión profunda de él para progresar en el campo del cálculo de probabilidades. En particular, la técnica de pedir a los estudiantes que discriminen secuencias aleatorias y no aleatorias para apreciar sus propiedades ha sido propuesta por estos autores.

Con relación al concepto de *espacio muestral*, se puede hallar una reseña de los estudios didácticos en el tema en Jones, et al. (2007, p. 920), donde se afirma que "*es un concepto fundamental en el proceso de matematización de los fenómenos aleatorios"*. Llama la atención un resultado de Jones, Langrall, Thornton y Mogill (1999), en el que cuando se pide a niños de 8-9 años que describan todos los posibles resultados de una experiencia simple, responden con *un sólo resultado* o una parte del espacio muestral, eliminando otros, porque han ocurrido en anteriores ocasiones. Es curioso que la conducta de responder con un solo resultado se mantenga con estudiantes con casi el doble de edad de los sujetos de Jones et al. (2007), como se verá más adelante.

Por otro lado, Jones, et al. (2007) señalan también que los conceptos de *variable aleatoria* y *distribución* se suelen introducir en el currículo de probabilidad del nivel Bachillerato; en particular, el concepto de distribución binomial. No obstante son escasos los trabajos de investigación didáctica que se refieren a la distribución binomial. Entre ellos, podemos mencionar los de Bill, et al. (2009) y Fischbein y Schnarch (1997) que se ubican en el Bachillerato, y los de Alvarado y Batanero (2007), Maxara y Biehler (2010) y Van Dooren, De Bock, Depaepe, Janssens y Verschaffel (2003) en el nivel universitario.

3. Marco conceptual

Se entiende por marco conceptual los principales conceptos que permiten entender el estudio. En el presente trabajo, se han elegido conceptos que se refieren a tres dimensiones: contenido, cognición y metodología. De la primera, se exponen contenidos relacionados con la distribución binomial; de la segunda la noción de razonamiento y de la tercera la taxonomía SOLO.

3.1. La distribución binomial

La *distribución binomial* es una de las distribuciones discretas más importantes en probabilidad, debido a que da lugar a gran número de problemas interesantes, a sus múltiples aplicaciones y a su relación con la distribución normal. Una distribución de Bernoulli con parámetro p, sólo tiene dos posibles resultados 0 y 1 (fracaso y éxito, respectivamente), donde p es la probabilidad de éxito y $(1 - p)$ la probabilidad de fracaso. Suponga que se repite n veces de manera independiente una experiencia modelada por una distribución de Bernoulli: entonces si se define la variable aleatoria "número de éxitos obtenidos en las n repeticiones", la distribución de probabilidades de esta variable aleatoria es una binomial, con parámetros n y p. El caso más simple de una distribución binomial (diferente de la Bernoulli) se presenta cuando estos parámetros son 2 y ½ respectivamente. Esta distribución es el modelo de probabilidad de un gran número de experiencias aleatorias, en particular, la de lanzar dos monedas (o equivalentemente dos veces una moneda) y observar el número de 'águilas' [nombre popular de una cara de las monedas mexicanas equivalente a "cruz"] que ocurren. La distribución $B(x, 2, ½)$ se representa fácilmente mediante una tabla o una gráfica, como se muestra en la Figura 1.

Representación tabular:

x	0	1	2
p	¼	½	¼

Representación gráfica:

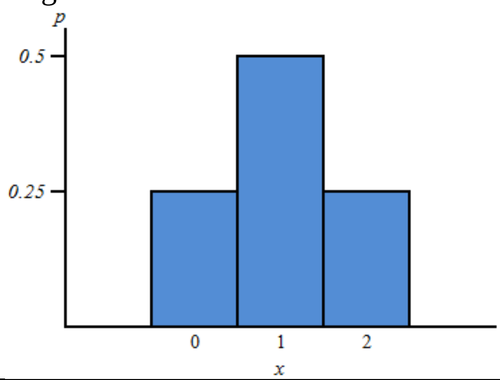

Figura 1. Representaciones de la distribución $B(x, 2, ½)$.

A lanzar un par de monedas legales y registrar el valor de la variable obtenida se le llama "observar la variable aleatoria". Una secuencia de resultados de observar muchas veces esta variable aleatoria binomial sería una secuencia desordenada de valores 0, 1 y 2, con frecuencias relativas cercanas a sus respectivas probabilidades. Formalmente, el *valor esperado* de cada valor de la variable en las N repeticiones es el producto Np_v donde p_v es la probabilidad de dicho valor. Así si $N = 20$, los valores esperados de 0, 1 y 2 son 5, 10 y 5 respectivamente. El *valor esperado* es una referencia importante de la distribución.

En la investigación en didáctica de la probabilidad y la estadística, por ejemplo, en Watson, Kelly, Callingham y Shaughnessy (2003), se ha introducido la técnica de hacer preguntas comenzando con la expresión ¿Qué esperas que ocurra si…? con la idea de explorar en qué medida los estudiantes tienen un sentido de la variabilidad presente en los experimentos. En contraste con una idea determinista (inadecuada) de la distribución, reducida sólo a *los valores esperados*; por ejemplo, preguntan ¿Con qué frecuencia crees que ocurra cada cara al lanzar 60 veces un dado? Las respuestas que no consideran la variación suelen decir "10 veces cada cara". Arteaga, Batanero, Cañadas y Contreras (2013) introdujeron y desarrollaron la técnica de pedir a profesores de primaria en formación que inventen una secuencia que consideren aleatoria y la comparen con otra real, para desarrollar su intuición del azar. En el Problema 1 de la presente investigación se han combinado aspectos de ambas técnicas.

3.2. Razonamiento en probabilidad

Se entiende por *razonamiento* al proceso de formular juicios o aseveraciones a partir de otras proposiciones ya conocidas, o de observaciones o reflexiones sobre un fenómeno. Asumimos que la gente se forma conceptos como consecuencia de su participación en actividades de buscar y pedir razones, es decir, gracias al *razonamiento* establecen y consolidan sus conceptos (Bakker & Derry, 2011).

En consecuencia, es conveniente proponer tareas o situaciones-problema que propicien el *razonamiento* de los estudiantes. Por ejemplo, en el campo de la probabilidad, tareas en situaciones de incertidumbre que logren captar su interés, lo enfrenten a algunas de las 'grandes ideas' de la probabilidad (Gal, 2005) y le exijan realizar razonamientos probabilísticos. Tales tareas deben involucrar a los estudiantes en procesos de estimación de la propensión de ocurrencia de eventos y de su predicción y representación, y no restringirse sólo a problemas de cálculo.

3.3. Taxonomía SOLO

Una forma de describir los razonamientos de los estudiantes ha sido mediante jerarquías de razonamiento. Por ejemplo, Jones, Langrall, Thornton y Mogill (1997) generaron y validaron un marco teórico, cuyo objetivo es caracterizar el razonamiento probabilístico de los estudiantes de primaria. El marco está basado en una síntesis de la literatura en probabilidad y fue apoyado por estudios empíricos de los investigadores con niños de 1º a 4º grado en un periodo de 2 años. Para cada constructo de probabilidad, se incluyen cuatro niveles de razonamiento, dentro de los modos icónico y simbólico-concreto del modelo de Biggs y Collis (Jones, et al., 2007, p. 930). En el presente estudio también se utiliza el modelo de Biggs y Collis (1991) para analizar las respuestas de los estudiantes y, con base en el análisis, describir sus razonamientos.

La *taxonomía SOLO* ha sido la base para construir jerarquías de comprensión o razonamiento en probabilidad y estadística. De manera similar a la teoría de J. Piaget sobre la etapas de desarrollo, el modelo neo-piagetiano SOLO de Biggs y Collis (1991) afirma que en el desarrollo intelectual de las personas se pueden identificar los siguientes *modos de representación*: 1) *Sensorio-motor*, comienza con el nacimiento; 2) *icónico* desde 18 meses, 3) *simbólico-concreto* a partir de los 6 años, 4) *formal* desde los 14 años y 5) *post-formal* de los 20 años para arriba.

Por supuesto que esos límites son un promedio, variando de acuerdo a las diferentes condiciones de las situaciones concretas. Los "*modos, entonces, son niveles de abstracción, progresando de acciones concretas a conceptos y principios abstractos, que forman la base de las etapas del desarrollo"* (Biggs & Collis, 1991, p. 62). Dentro de cada modo y, con relación a algún tipo de conocimiento, los aprendices despliegan secuencias, o ciclos de aprendizaje, que progresan de la incompetencia a la pericia. Dicha secuencia tiene un crecimiento jerárquico que se refleja en la complejidad estructural de las respuestas. "*Esta jerarquía nos dice qué tanto el aprendizaje ha progresado hacia la pericia [...] y puede por lo tanto ser utilizada para clasificar los resultados del aprendizaje dentro de cualquier modo"* (p. 64) por esta razón se le llama *Taxonomía* SOLO. Ésta postula, además de los modos, ciclos de aprendizaje con cinco niveles:

1. En el nivel Pre-estructural (P) se observa en las respuestas que se utilizaron sólo aspectos irrelevantes a la tarea.
2. En el nivel Uniestructural (U) las respuestas reflejan sólo un aspecto relevante y por tanto, no se llega a la respuesta correcta.
3. En el nivel Multiestructural (M) se identifica en las respuestas varios aspectos relevantes a la tarea, pero no de manera integrada.
4. En el nivel relacional (R) las respuestas indican que se identifican los aspectos relevantes a la tarea y se integran de manera conveniente.
5. Finalmente, en el nivel Abstracto Extendido (AE) las respuestas reflejarían que el estudiante fue capaz de generalizar la estructura, para incorporan rasgos abstractos que representan el pensamiento en un modo superior.

Un ejemplo importante de aplicación de la Taxonomía SOLO con relación a la didáctica de la probabilidad lo constituye el trabajo de Watson (2006, p. 13), quien comenta que "*es un instrumento para clasificar las respuestas de los estudiantes a alguna tarea; enfatiza lo que es observado en dichas respuestas y no en lo que el observador cree que los estudiantes pudieran haber entendido".* La jerarquía de razonamiento para el pensamiento probabilístico propuesta por Jones, et al. (1997) – de acuerdo al autor – es igualmente consistente con la taxonomía SOLO.

4. Metodología

Se diseñó un cuestionario con preguntas a partir de una situación binomial, en la que se intenta propiciar que los estudiantes consideren aspectos de la aleatoriedad, y de la definición clásica y frecuencial de la probabilidad de manera espontánea, es decir, no precedidas por una instrucción sobre esos temas.

Se exploran las respuestas de estudiantes de dos niveles escolares, secundaria y Bachillerato para poder detectar algunos rasgos de los razonamientos que se presentan en este periodo de transición. Conviene aclarar que suponemos que los estudiantes, tanto de secundaria como de Bachillerato, se ubican en el *modo de representación* *formal*. De acuerdo al modelo de Biggs y Collis (1991, p. 63), los sujetos en este modo de representación son capaces de implicarse con sistemas abstractos dentro de los cuales se organizan muchos elementos y pueden ser usados para generar hipótesis. Por esta razón, se asume que dentro de este modo se construye la noción de distribución binomial que estamos explorando.

4.1. Participantes

Como se ha dicho, en este informe participan estudiantes de dos grupos; uno pertenece al sistema de secundaria y otra a un sistema de Bachillerato de la Universidad Nacional Autónoma de México (UNAM), ambos ubicados en la Ciudad de México. El primero está formado por treinta y cuatro estudiantes de tercer grado de secundaria (14-15 años) y el segundo por cuarenta y seis estudiantes del primer grado de Bachillerato (15-16 años). El profesor titular del grupo de Bachillerato y dos investigadores, quienes son los autores del presente artículo, más el maestro de los estudiantes de secundaria, completan el equipo de investigadores. Cuando se realizó la actividad que aquí se reporta, los estudiantes de ambos grupos no habían recibido ningún tipo de enseñanza formal de probabilidad y estadística.

4.2. Instrumento

Se elaboró una actividad de 13 preguntas con base en una situación-problema llamada *"¡A la Suerte!"* que se expone en la Figura 2. A continuación se exponen tres de las preguntas de la actividad y se añaden comentarios sobre el propósito de cada una (en la actividad inicial se hicieron 13 preguntas, de las que acá se analizan solo tres). Algunas de las preguntas fueron ligeramente diferentes en los dos grupos.

¡A LA SUERTE!

La familia Pérez, está formada por el señor Carlos, su esposa Ana y su hijo Beto. Todas las noches, por lo general, después de cenar se reúne la familia a ver la televisión, pero nunca están de acuerdo para ver un mismo programa. Al señor Carlos le gusta ver programas deportivos, a la señora Ana las películas y a Beto las caricaturas.

Como sólo hay un televisor en la casa, lo más sencillo sería que se turnaran el control de la televisión diariamente, pero Beto les propone a sus padres algo más divertido: "¡A la suerte!".

Propone rifar el control lanzando dos monedas de la siguiente manera: Si no sale ninguna águila en los dos volados, gana la Sra. Ana; si sale exactamente un águila, gana el niño Beto; y si salen dos águilas, gana el Sr. Carlos. A los padres les parece justo y aceptan su propuesta

Figura 2. Actividad llamada "¡A la Suerte!".

La pregunta 1 (Figura 3) tiene el propósito de explorar la forma en que los estudiantes perciben la aleatoriedad en una secuencia de tres resultados posibles. Se espera que propongan secuencias desordenadas que reflejen la aleatoriedad y cuyas frecuencias prefiguren la distribución que atribuyen, consciente o inconscientemente, a la variable persona ganadora: Ana, Beto, Carlos.

1. Escribe la secuencia de resultados que crees que pueda ocurrir al jugar durante los primeros 20 días. (Marca una **X** por cada resultado en cada uno de los 20 días).

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**T**
Ana																					
Beto																					
Carlos																					

Figura 3. Pregunta 1.

Cuestionario de Jaimes (2011), aplicado a estudiantes de Secundaria:
2. ¿Cuántas águilas pueden salir en dos volados?

Cuestionario de Medina (2013), aplicado a estudiantes de Bachillerato:
2. ¿Cuáles son los posibles resultados del número de águilas que se obtienen en un sorteo de los dos volados?

Figura 4. Pregunta 2.

En la pregunta 2 (Figura 4) se pide el recorrido de la variable aleatoria: {0, 1, 2} y se espera su descripción en forma de lista o descrita verbalmente; es aceptable que exhiban los elementos del espacio muestral {AA, AS, SA, SS} ya que este conjunto contiene, implícitamente, los valores de la variable y favorece las condiciones para la percepción o cálculo de la probabilidad. La formulación de la pregunta es diferente en cada grupo, pues se hicieron en momentos distintos. Después de realizar el estudio de Jaimes (2011) con estudiantes de secundaria, se decidió mejorar la formulación de la pregunta para el estudio en Bachillerato (Medina, 2013).

3. ¿Crees que después de muchos días, todos los miembros de la familia tendrían el mismo número de veces el control del televisor? Justifica tu respuesta.

Figura 5. Pregunta 3.

El propósito de la pregunta 3 (Figura 5) es explorar las razones por las que el estudiante cree que algún miembro de la familia tiene ventaja o no. Se esperaba que lo dedujeran a partir del examen del espacio muestral. En la pregunta 1 ya se puede percibir quienes atribuyen más probabilidad a algún miembro de la familia, el cual debe ser Beto, pero con este problema se confirmaría si lo aplican al comportamiento a largo plazo de las frecuencias.

4.3. Procedimiento

El docente presenta a sus estudiantes la situación-problema, les pide leerla en forma individual y grupal, les permite que realicen algunos ensayos para propiciar una familiarización con la situación aleatoria y puedan opinar, hacer predicciones, argumentar y proponer ideas en el cuestionario. Finalmente, los estudiantes completan el cuestionario. La aplicación duró 1 hora aproximadamente.

Procedimiento de análisis. Las respuestas obtenidas se organizan de acuerdo a su complejidad estructural, bajo el supuesto de que entre mayor es ésta, mejor la calidad de la respuesta. Para hacerlo se ha tenido como guía la taxonomía SOLO (Biggs & Collis, 1982, 1991). Mediante un análisis *a priori* de la tarea, se identifican sus componentes, es decir, los diferentes conceptos y procedimientos que convenientemente relacionados ofrecen una solución y respuesta satisfactoria. Con base en la identificación de tales componentes, se definen niveles crecientes de complejidad estructural, intentando aplicar el esquema expuesto en el Marco Conceptual. No obstante, a la hora de hacerlo se hace necesario realizar algunas modificaciones a dicho esquema por las siguientes razones: a) emergen componentes no contempladas al principio y que conviene incorporarlas; b) hay componentes previstas que no se aplican a las respuestas obtenidas, c) se incorporan como elementos importantes en la clasificación los errores presentes en las respuestas y, d) no es posible definir componentes independientes. Con base en estos considerandos, se definen los niveles para cada pregunta de la siguiente manera.

Niveles para la pregunta 1. El primer aspecto o componente que se debe reflejar en la respuesta es que se entienda lo que se pide, en particular, que se señale un sólo ganador por sorteo de manera que la suma total sea 20. Si una respuesta no cumple con esta componente se clasifica en pre-estructural, pues es imposible que se satisfagan las otras si no se entiende lo que se pide. Una segunda componente es que la secuencia que proponga el estudiante tenga características de una secuencia aleatoria, es decir, que no tenga patrones repetidos y claramente perceptibles, o bien, tenga rachas convenientes: una racha de longitud 3 o al menos dos rachas de longitud 2. Mediante una simulación de 1000 secuencias de tamaño 20, se estimó que $P(A) \approx 0.90$; $P(B) \approx 0.80$; y además, $P(A \cup B) \approx 0.93$; donde A = "Se obtiene una racha de tamaño 3", y B = "Se obtiene al menos dos rachas de tamaño 2". Una tercera componente es que en la columna "Total" se refleje la distribución binomial, es decir, que se atribuya mayor frecuencia de éxito al segundo jugador: Beto. Si se cumple la primera componente pero no las demás, la respuesta es Uniestructural; si se satisface además otra u otras componentes, pero sin integrarlas todas, se clasifica en Multiestructural; finalmente, si todas se consideran de manera integral, la respuesta se clasifica en Relacional.

Niveles para la pregunta 2. Las respuestas que no producen sentido desde el punto de vista de lo que se pide, se clasifican en Preestructural. Si la respuesta consiste en un sólo elemento del recorrido de la variable aleatoria o en un sólo elemento del espacio muestral, se clasifica en Uniestructural. En caso de que la respuesta consista en dos elementos del recorrido de la variable aleatoria, o bien, en dos o tres elementos del espacio muestral, se considera Multiestructural y, finalmente, si la respuesta ofrece los tres elementos de la variable aleatoria o los cuatro elementos del espacio muestral se clasifica en Relacional. El siguiente análisis apoya dicha categorización: los estudiantes que dan respuestas en el nivel Uniestructural pueden estar pensando que la unidad de análisis es sólo una experiencia, o bien, que lo que se les pregunta es ¿Cuál es el máximo de águilas que pueden ocurrir?, en este caso responden 2. En ambos casos, ven sólo un resultado. En cambio, los estudiantes que responden con dos valores de la variable aleatoria o con dos o tres elementos del espacio muestral, ya piensan en varias posibilidades de ocurrencia, pero lo hacen de manera incompleta o excluyen el resultado en que "no ocurre águila". Finalmente, los que responden con todos los valores de la variable y/o todos los elementos del espacio muestral tienen una noción básica del recorrido de la variable aleatoria.

Niveles para la pregunta 3. La formulación de esta pregunta provoca, en primer lugar, las respuestas “Sí” o “No” y en seguida una argumentación para apoyar la elección hecha. Cuando la respuesta es “Sí” se clasifica en Preestructural ya que el estudiante piensa que cada participante ganará al final de muchas repeticiones el mismo número de veces, posiblemente por el sesgo de equiprobabilidad. En caso de que respondan “No”, se abren varias posibilidades. Las respuestas que dicen “No” sin ninguna explicación adicional, o cuyas razones son diferentes a la de evaluar que Beto tiene mayor ventaja, se clasifican en Uniestructural, ya que reconocen que no tiene por qué haber la tendencia de que todos ganen con la misma frecuencia. Las respuestas que dicen “No” pero la explicación se basa en la variabilidad o identifican el valor 1 (Beto) como más probable, sin más explicación, se clasifican en Multiestructural. Las explicaciones basadas en la variabilidad son aquellas que dicen que “No” porque perciben es muy poco probable los resultados se distribuyan de manera equitativa; ésta es una intuición adecuada. Finalmente, las respuestas que ofrecen explicaciones que se basan en el conocimiento de la distribución binomial se clasifican en Relacional.

5. Resultados

En las tablas 2, 3 y 4 se muestran las frecuencias de respuesta por nivel estructural a las preguntas 1, 2 y 3, respectivamente. En seguida se exponen de manera resumida los resultados obtenidos por cada pregunta.

5.1 Pregunta 1. Características de la secuencia de resultados prevista

En esta pregunta analizamos los resultados previstos. Con relación a señalar un sólo ganador por sorteo, de manera que la suma total sea 20, en secundaria aún se presentan 7 casos en que no entienden lo que se les pide, mientras que en Bachillerato sólo 1; el complemento está formado por respuestas que satisfacen esta componente. Con relación a la segunda componente, la mayoría propone secuencias sin patrones y con rachas adecuadas. Las frecuencias en las que se proponen secuencias de rachas de diferentes longitudes se muestran en la Tabla 1. No se observan grandes diferencias en los dos grupos.

Tabla 1. *Frecuencias de secuencias con diferentes rachas en la pregunta 1*

Longitud de la racha	Secundaria		Bachillerato	
(Nº. de rachas)	Frecuencia	Porcentaje[1]	Frecuencia	Porcentaje[1]
2 (al menos dos)	20	91	30	94
3 (al menos una)	9	41	14	44
4 (al menos una)	3	14	5	16

[1] La columna no suma 100% por que los eventos no son excluyentes.

Finalmente, la componente en la que hay mayor diferencia entre los estudiantes de secundaria respecto a los de Bachillerato es la referente a la distribución binomial. Las propuestas de frecuencias de los estudiantes de secundaria tienden a ser uniformes, pues aproximadamente se espera que cada jugador gane igual número de veces (Ver ejemplo en Figura 6), mientras que las frecuencias de los estudiantes de Bachillerato ya tienden a los valores esperados derivados de la distribución, donde el suceso una sola águila se ve con doble probabilidad que los otros dos (Ver ejemplo en Figura 7)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	T
Ana		X		X							X	X		X				X		X	7
Beto	X					X		X	X				X			X			X		7
Carlos			X		X		X			X					X		X				6

Figura 6. Respuesta del estudiante ES6.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	T
Ana			X				X			X		X							X		5
Beto	X			X	X				X				X	X	X		X	X		X	10
Carlos		X				X		X			X					X					5

Figura 7. Respuesta del estudiante EB30.

En resumen, tanto los estudiantes de secundaria como de Bachillerato tienen intuiciones adecuadas relacionadas con el desorden y la longitud de las rachas, pero los de secundaria no perciben la distribución binomial que gobierna los sorteos, dejándose llevar por el sesgo de equiprobabilidad. Al parecer, los estudiantes de Bachillerato espontáneamente pueden analizar la situación y descubrir que el segundo jugador tiene el doble de ventaja que cualquiera de los otros dos. En la Tabla 2 se pueden ver las frecuencias en las que se clasificaron las respuestas a la pregunta 1 en los diferentes niveles en la taxonomía SOLO.

Tabla 2. *Frecuencias de respuestas a la pregunta 1 por nivel estructural.*

Secundaria			Bachillerato		
Nivel de razonamiento	Frecuencia	Porcentaje	Nivel de razonamiento	Frecuencia	Porcentaje
0. Preestructural	7	21	0. Preestructural	1	3
1. Uniestructural	3	9	1 .Uniestructural	8	18
2. Multiestructural	14	41	2 .Multiestructural	8	18
3. Relacional	10	29	3. Relacional	27	61
Total	34	100	Total	44	100
$\mu_{1S} = 1.79$			$\mu_{1B} = 2.38$		

5.2. Pregunta 2. Construcción del espacio muestral

Estudiamos en esta pregunta la comprensión del espacio muestral. En la Tabla 3 se muestra cómo se distribuyeron las frecuencias de la pregunta por nivel estructural. Llama la atención que en esta pregunta, a diferencia de lo que ocurre con las otras, haya más respuestas de estudiantes de Bachillerato clasificadas en Preestructural, que respuestas de los estudiantes de secundaria. Es posible que la formulación de la pregunta a los de Bachillerato hiciera más difícil entender lo que se pedía (describir el recorrido de la variable o el espacio muestral), aunque el propósito al reformularla fue hacerla más clara, pues en la versión original que se les hizo a los estudiantes de secundaria provocaba confusión; sin embargo, no tenemos más evidencia para apoyar esta hipótesis. Otra explicación es que los estudiantes de Bachillerato posiblemente creen que la pregunta pide más de lo que realmente se solicita, y responden a una pregunta más sofisticada, por ejemplo, EB11: *Pues hay una probabilidad del 50% porque cada moneda tiene dos caras.*

Tabla 3. *Frecuencias de respuesta a la pregunta 2 por nivel estructural.*

Secundaria			Bachillerato		
Nivel de razonamiento	Frecuencia	Porcentaje	Nivel de razonamiento	Frecuencia	Porcentaje
0 Preestructural	6	18	0 Preestructural	14	33
1 Uniestructural	27	79	1 Uniestructural	17	39
2 Multiestructural	1	3	2 Multiestructural	3	7
3 Relacional	0	0	3 Relacional	9	21
Total	34	100	Total	43	100
$\mu_{2S} = 0.88$			$\mu_{2B} = 1.16$		

En ambos niveles son muy frecuentes las respuestas clasificadas en el nivel Uniestructural, por ejemplo: ES4: *2* o EB37: *Que salga solamente 1 águila.* Mientras que en el nivel Multiestructural hay sólo 1 de secundaria y 3 de Bachillerato, por ejemplo: ES2: *2 o 1* o EB9: *Que uno salga águila y el otro sol, que los dos salgan águila o que los dos salgan sol*. Es interesante notar que no hubo respuestas de estudiantes de secundaria en el nivel relacional, en contraste, 9 del Bachillerato se clasificaron en este nivel, como EB21: *Pues puede salir sólo estas combinaciones águila-sol, águila-águila, sol-águila, sol-sol, 2 de 4.* Otro rasgo que vale la pena considerar es que sólo los estudiantes de Bachillerato consideran espontáneamente los elementos del espacio muestral *AA, SA, AS, SS,* mientras que los de secundaria sólo consideran valores de la variable 0, 1 y 2.

Es algo sorprendente que la descripción de la variable aleatoria o del espacio muestral sea tan difícil, pues en realidad es algo muy simple. Sin duda influye la forma en que se formula la pregunta, pero sospechamos que no es sólo eso o que la dificultad en formular la pregunta refleja sutiles dificultades conceptuales. La conducta de responder con un sólo elemento del espacio muestral o del recorrido de la variable ya había sido identificada por Jones et al. (1999) en niños de entre 10 y 12 años de edad. Estas respuestas, clasificadas en Preestructural, se explicarían si se supone que el estudiante imagina un sólo sorteo de la experiencia y responde lo que podría ser el resultado de un sorteo. Sería un razonamiento similar al enfoque del resultado aislado de Konold (1991), en el que los estudiantes, ante ciertos problemas de probabilidad, quieren predecir lo que va a pasar en una sola prueba.

En el caso de las respuestas clasificadas en nivel Multiestructural, la dificultad principal es imaginar los resultados "ausencia de águila" cuando la variable dice "el número de águilas", esta formulación parece llamar la atención sobre los casos en los que se presenta efectivamente al menos un águila. Sólo en los estudiantes de Bachillerato se presentan manifestaciones espontáneas de un pensamiento combinatorio, con el que logran enlistar todas las posibilidades.

5.3. Pregunta 3. Distribución esperada

En la pregunta 3 se pide a los estudiantes que respondan si están de acuerdo en que a la larga cada participante tendrá el televisor el mismo número de veces y lo justifiquen. En la Tabla 4 se pueden observar las frecuencias en las que se clasificaron las respuestas en los diferentes niveles de la taxonomía SOLO.

La frecuencia de respuestas en el nivel Preestructural es mayor, como se espera, en los estudiantes de secundaria y, en general, las respuestas clasificadas en este nivel están influidas por el sesgo de equiprobabilidad (Lecoutre, 1992), por el que se espera que todos los resultados de un experimento aleatorio sean equiprobables. Por ejemplo, encontramos este razonamiento en ES24: *Si porque todos tienen las mismas posibilidades* o EB11: *Si porque todos tienen la misma probabilidad de ganar*. Un caso especial es el estudiante EB41, quien responde: *podría pasar que sí, pero dentro de mucho tiempo*, es decir, admite que a la larga los tres resultados son igualmente posibles. Sin embargo, en la pregunta 1 había propuesto una secuencia de resultados que llevan a la distribución de frecuencias 5, 10, 5. Es decir, identificó la distribución binomial.

Tabla 4. *Frecuencias de respuesta a la pregunta 3 por nivel estructural.*

Secundaria			Bachillerato		
Nivel de razonamiento	Frecuencia	Porcentaje	Nivel de razonamiento	Frecuencia	Porcentaje
0 Preestructural	14	41	0 Preestructural	5	12
1 Uniestructural	20	59	1 Uniestructural	22	52
2 Multiestructural	0	0	2 Multiestructural	13	31
3 Relacional	0	0	3 Relacional	2	5
Total	34	100	Total	42	100
$\mu_{2S} = 0.59$			$\mu_{2B} = 1.29$		

La mayoría de las respuestas a esta preguntan se concentran en el nivel Uniestructural, tanto en secundaria cómo en Bachillerato; ya que reconocen que no tiene por qué haber la tendencia de que todos ganen con la misma frecuencia, pero por razones diferentes a la de evaluar que Beto tiene ventaja: ES13: *No, porque no siempre gana uno*, ES18: *No, porque no siempre saldrán los mismos resultados*, EB1: *No, sería muy mínima la diferencia que habría pero serían diferentes los días*.

Ningún estudiante de secundaria argumentó con base en la identificación de que Beto tiene mayor probabilidad; en cambio, en Bachillerato varios lo hicieron, situándose en nivel Multiestructural: EB25: *No, Beto siempre lo tendrá más veces porque tiene más probabilidad de ganar*. E incluso alguno llegó al relacional, estimando correctamente la probabilidad de cada caso.

5.3. Síntesis de respuestas

En la Tabla 5, se comparan los índices de respuestas para cada pregunta y por grupo. Éstos se calculan asignando los valores 0, 1, 2 y 3 respectivamente a cada nivel estructural (Prestructural: 0, Uniestructural: 1, etc.) y calculando la media ponderada de las frecuencias; en consecuencia, estos índice son números entre 0 y 3. En dicha tabla, se observa que el orden de dificultad de las preguntas, para los estudiantes de secundaria progresa en el mismo orden en que están numeradas, pues los índices disminuyen. En cambio, para los estudiantes de Bachillerato la pregunta 2 resulta la más difícil.

Tabla 5. *Comparación de índices de respuestas.*

	Secundaria	Bachillerato	Diferencia de medias
Pregunta 1	1.79	2.38	0.59
Pregunta 2	0.88	1.16	0.28
Pregunta 3	0.59	1.29	0.70

Globalmente, el desempeño de los estudiantes de Bachillerato (15-16) es mejor que el de los estudiantes de secundaria (13-14), pues en todas las preguntas la diferencia de medias es positiva. Las diferencias de los índices son indicadores de dicho progreso, donde se nota que en la pregunta 2 hay un avance algo menor que el experimentado en las otras dos.

La pregunta 1, en la que se pide inventar una posible secuencia de 20 sorteos, tiene índices de respuestas mayores con respecto al de las otras dos: 1.79 en secundaria y 2.38 en Bachillerato (Tabla 5); por tanto fue la más sencillas de las propuestas.

En la pregunta 2 se pide describir los valores de la variable aleatoria y, aunque no se preguntó sobre el espacio muestral, se consideran convenientes las respuestas que lo describen. Se puede observar que el índice de respuesta para secundaria es 0.88 y para Bachillerato 1.16, este es el más bajo con relación a los otros dos índices de Bachillerato (Tabla 5).

El índice de respuestas para nivel secundaria en la pregunta 3 (predecir la distribución) es el más bajo (0.59) con relación al de las otras preguntas (Tabla 5).

6. Discusión y conclusiones

En este estudio se describen las pautas en los razonamientos probabilísticos de los estudiantes de finales de secundaria y principios de Bachillerato que permiten describir niveles estructurales en sus respuestas a tres preguntas de probabilidad relacionadas con la noción de distribución binomial. En seguida se destacan, para cada pregunta, los aspectos más importantes que se observaron en el análisis de sus respuestas.

Tanto los estudiantes de secundaria como los de Bachillerato componen secuencias que tienen dos características de la aleatoriedad: no presentan patrones repetidos y tienen rachas de longitud 2 o más. Los estudiantes de Bachillerato ofrecen respuestas más avanzadas pues además consideran la distribución, es decir, perciben que el segundo jugador (Beto) tiene el doble de probabilidad que la probabilidad de los otros dos valores de la variable (Ana y Carlos). Una explicación posible podría ser asumir que los estudiantes de Bachillerato tienen un pensamiento combinatorio más avanzado que los de secundaria.

Una parte significativa de los estudiantes examinados responden a la segunda pregunta con un sólo valor del recorrido de la variable o del espacio muestral, en lugar de dar la descripción completa de alguno de ellos. Los que responden con más de un valor, pero sin dar todo el recorrido o el espacio, suelen omitir el evento "ausencia de águilas". En esta pregunta es menor la diferencia, con relación a las otras dos preguntas, entre el índice de respuesta de los de Bachillerato y de secundaria; esto a pesar de la reformulación que se hizo a la pregunta aplicada a los de secundaria antes de aplicarla a los de Bachillerato.

Aún se encuentra un buen porcentaje de estudiantes cuyas respuestas a la tercera pregunta indican que a la larga los tres participantes van a tener el control aproximadamente el mismo número de veces; es posible que estas respuestas se deban al sesgo de equiprobabilidad (Lecoutre, 1992). En los casos en que niegan que va a haber un equilibrio, su argumento no se basa en que se hayan dado cuenta que la distribución tiene más probabilidad para Beto, sino más bien por un sentido de variabilidad: "*sería mucha casualidad que a los tres les hubiera tocado el mismo número de veces*".

Es algo sorprendente la dificultad de los estudiantes para describir el recorrido de la variable (o el espacio muestral) y que un buen porcentaje responde con uno o dos elementos pero no todo el recorrido. Es posible que cuando se les pregunta ¿Cuántas águilas pueden salir en dos lanzamientos de la moneda? Los estudiantes asumen que la respuesta debe ser un valor individual, ya que varios responden "2" quizá pensando en el máximo número de águilas; algunas respuestas proponen otro valor aislado de la variable. No hay una mejora importante al reformular la pregunta para los estudiantes de Bachillerato. La pregunta: ¿Cuáles son los posibles resultados del número de águilas que se obtienen en un sorteo de los dos volados? también produce respuestas consistentes en valores aislados. Esta conducta ya la había identificado Jones, et al. (1999) con estudiantes de menor edad que los nuestros. Por otro lado, al estudiar las respuestas a preguntas relacionadas con la variación estadística, Shaughnessy (1997) comenta que los estudiantes no suelen responder con rangos de valores sino con valores únicos; y nuestros resultados lo apoyan. En consecuencia, en la enseñanza conviene no asumir que el concepto de espacio muestral queda firmemente establecido con explicaciones o con una sola actividad.

Menos sorprendente pero igualmente importante es la persistencia del sesgo de equiprobabilidad descrito por Lecoutre (1992). Éste se produce al no percibir la relación de la variable con el espacio muestral ya que tratan al recorrido de la variable como si fuera el espacio muestral (equiprobable). También llama la atención que un estudiante de Bachillerato haya proporcionado frecuencias en la pregunta 1 de acuerdo a la distribución binomial y, sin embargo, cayera en el sesgo de equiprobabilidad en la pregunta 3.

Con relación a la pregunta 1, las observaciones indican que para muestras de tamaño 20 las intuiciones de los estudiantes sobre la aleatoriedad son adecuadas ya que proponen secuencias que no presentan patrones repetidos y con rachas de longitud razonables. Las pautas identificadas en este estudio pueden contribuir al diseño de actividades para ofrecer oportunidades de aprendizaje de la distribución binomial para estudiantes del nivel explorado. En particular, se recomienda evitar la tendencia a refugiarse sólo en el cálculo, evitando los difíciles conceptos de aleatoriedad y variabilidad. Por el contrario, puede ser productivo que en las actividades los estudiantes hagan observaciones de la variable aleatoria y reflexionen sobre lo que puede ocurrir cuando repiten muchas veces el experimento, además de los cálculos ya conocidos.

Reconocimiento

Proyecto EDU2013-41141-P (Ministerio de Economía y Competitividad).

Referencias bibliográficas

Alvarado, H., & Batanero, C. (2007). Dificultades de comprensión de la aproximación normal a la distribución binomial. *Números, 67*, 1-7.

Arteaga, P., Batanero, C., Cañadas, G. R., & Contreras, J. M. (2013). Prospective primary school teachers' errors in building statistical graphs. En B. Ubuz, Ç. Haser, & M. Mariotti (Eds,). *Proceedings of the Eight Congress of the European Society for Research in Mathematics Education* (pp. 745-755). Antalya, Turkey: ERME.

Bakker, A., & Derry, J. (2011). Lessons from inferentialism for statistics education. *Mathematical Thinking and Learning*, *13*(1-2), 15-26.

Batanero, C., Henry, M., & Parzysz, (2005). The nature of chance and probability. En G. A. Jones (Ed.), *Exploring probability in school. Challenges for teaching and learning* (pp.20-42). New York: Springer.

Batanero, C., Serrano, L. & Green, D. R. (1998). Randomness, its meanings and implications for teaching probability. *International Journal of Mathematics Education in Science and Technology, 29* (1), 113-123.

Biggs, J. B., & Collis, K. (1982). *Evaluating the quality of learning. The SOLO taxonomy*. New York: Academic Press.

Biggs, J. B., & Collis, K .F. (1991). Multimodal learning and the quality of intelligence behavior. En H.A. Rowe (Ed.), *Intelligence: Reconceptualization and measurement* (pp. 57-76). Hillsdale, NJ: Lawrence Erlbaum.

Bill, A., Watson, J., & Gayton, P. (2009). Guessing answers to pass a 5-item true false test: solving a binomial problem in three different ways. *Proceedings of the 32nd annual conference of the Mathematics Education Research Group of Australasia, 1* (pp. 57-64). Tasmania: MERGA.

Borovcnik, M. (2012). Multiple perspectives on the concept of conditional probability. *Avances de Investigación en Educación Matemática*, *2*, 5-27.

Borovcnik, M., & Bentz, H. J. (1991). Empirical research in understanding probability. En R. Kapadia, M. Borovcnic (Eds.), *Chance encounters: Probability in education* (pp. 73-105). Dordrecht: Kluwer.

Fischbein, E., & Schnarch, D. (1997). The evolution with age of probabilistic, intuitively based misconceptions. *Journal for Research in Mathematics Education, 28*(1), 96-105.

Gal, I. (2005). Towards "probability literacy" for all citizens: Building blocks and instructional dilemmas. En G.A. Jones (Ed.), *Exploring probability in school. Challenges for teaching and learning* (pp. 39-63). New York: Springer.

Garfield, J. (1995). How students learn statistics. *International Statistical Review, 63*(1), 25-34.

Hacking, I. (1975). *The emergence of probability*. Cambridge: Cambridge University Press.

Jaimes, E. (2011). *Niveles de razonamiento probabilístico con énfasis en la noción de distribución de estudiantes de secundaria en tareas de experimentación y simulación computacional*. (Tesis de Maestría no publicada). Cinvestav – IPN, México D.F.

Jones, G.A., Langrall, C. W., & Mooney, E. S. (2007). Research in probability. En F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 909-955). Charlotte, NC, USA: Information Age-NCTM.

Jones, G. A., Langrall, C. W., Thornton, C. A., & Mogill, A. T. (1997). Framework for assessing and nurturing young children's thinking in probability. *Educational Studies in Mathematics, 32,* 101-125.

Jones, G. A., Langrall, C. W., Thornton, C. A., & Mogill, A. T. (1999). Students' probabilistic thinking in instruction. *Journal for Research in Mathematics Education, 30,* 487-519.

Konold, C. (1991). Understanding students' beliefs about probability. En E. Von Glasersfeld (Ed.), *Radical constructivism in mathematics education* (pp. 139-156). The Netherlands: Kluwer.

Lecoutre, M. P. (1992). Cognitive models and problem spaces in "purely random" situations. *Educational Studies in Mathematics, 23,* 557-568.

Maxara, C., & Biehler, R. (2010). Students' understanding and reasoning about sample size and the law of large numbers after a computer-intensive introductory course on stochastics. En C. Reading (Ed.), *Proceedings of the Eighth International Conference on Teaching Statistics* Ljubljana, Slovenia: Voorburg, The Netherlands: International Statistical Institute

Medina, M. (2013). *Niveles de razonamiento probabilístico ante una situación binomial por estudiantes de bachillerato*. (Tesis de Maestría no publicada), Cinvestav–IPN: México D.F.

Shaughnessy, J.M. (1992). Research in probability and statistics: Reflections and directions. En D.A. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning* (pp. 465-494). Reston, Va., USA: National Council of Teachers of Mathematics.

Shaughnessy, J. (1997). Missed opportunities in research on the teaching and learning of data and chance. En F. Biddulph, & K. Carr (Eds), *People in mathematics education* (pp. 6-22). Rotorua, New Zeland: MERGA.

Van Dooren, W., De Bock, D., Depaepe, F., Janssens, D., & Verschaffel, L. (2003). The illusion of linearity: expanding the evidence towards probabilistic reasoning. *Educational Studies in Mathematics*, *53*(2), 113-138.

Von Glasersfeld, E. (1990). An exposition of constructivism: Why some like it radical. *Journal for Research in Mathematics Education. Monograph No. 4.*

Watson, J. (2006). *Statistical literacy at school*. Mahwah, NJ: Lawrence Erlbaum.

Watson, J. M., Kelly, B. A., Callingham, R. A., & Shaughnessy, J. M. (2003). The measurement of school students' understanding of statistical variation. *International Journal of Mathematical Education in Science and Technology, 34*(1), 1-29.

Referencias a los autores

Jaime I. García, Departamento de Matemática Educativa, Cinvestav – IPN (México). jaime_garcia10@yahoo.com.mx

Miguel Medina, Departamento de Matemática Educativa, Cinvestav – IPN (México). ingcivi_6@yahoo.com.mx

Ernesto Sánchez, Departamento de Matemática Educativa, Cinvestav – IPN (México). esanchez@cinvestav.mx, esanchez0155@gmail.com

Levels of reasoning and abstraction of middle and high school students in a problem-situation of probability

Jaime I. García[1], Cinvestav – IPN (México)

Miguel Medina[2], Cinvestav – IPN (México)

Ernesto Sánchez[3], Cinvestav – IPN (México)

In this research, the way in which middle school and high school students solve a task that involves some basic notions of probability theory is explored. The aim is to identify the characteristic features of the levels of reasoning that middle and high school students exhibit when they face a probability situation-problem. A questionnaire with 13 questions about a simple binomial situation (n=2) was given to 34 students in the last grade of middle school (14-15 years) and 46 students in the first grade of high school (15-16 years). Neither of both groups had studied probability. In this paper we only analyze questions 1, 2 and 3. These questions, and the whole questionnaire, involve elementary aspects of randomness, sample space, random variable, and binomial distribution. The answers to these questions are analyzed to describe the levels of structural complexity using the SOLO taxonomy.

In the responses to Question 1 we observed that the students in both groups have similar suitable intuitions related to the disorder of sequences of results and the length of runs. However, middle school students do not perceive the distribution governing the experiment, responding according to the equiprobability bias, while many high school students identified the correct distribution. The responses to Question 2 show where we ask for the description of the random variable or the sample space was difficult to understand by the students. Many students in both groups answered with a value of the random variable or with an element of the sample space. This behavior had already been identified by Jones et al. (1999) in children between the ages of 10 and 12. In relation to Question 3, many middle school students accepted (inappropriately) that in the long run, the number of times each participant will win should be balanced, while the high school students rejected this idea. However, the argument of them was based on the variability of results and not on the distribution. In general, high school students showed a higher level of reasoning than those of middle school students.