Revista Integración
Escuela de Matemáticas
Universidad Industrial de Santander
Vol. 31, No. 1, 2013, pág. 69–86

Efectos de la simulación en la comprensión de la ley de los grandes números

Gabriel Yáñez Canal[a,*], Édgar Jaimes[b]

[a]Universidad Industrial de Santander, Escuela de Matemáticas, A.A. 678, Bucaramanga, Santander, Colombia.
[b]Instituto Técnico Industrial, Puente Nacional, Colombia.

Resumen. En este trabajo se presentan algunos de los resultados obtenidos en una investigación que indagaba por los procesos de razonamiento probabilístico de doce estudiantes entre 12 y 15 años de un colegio público en un municipio en el departamento de Santander (Colombia), alrededor de la Ley de los Grandes Números. Después de vivir una experiencia directa, los estudiantes realizaron simulaciones utilizando *Probability Explorer*. Los resultados sugieren que para tomar conciencia de la variabilidad en el corto plazo y de la estabilidad a largo plazo de las frecuencias relativas deben realizarse actividades bidireccionales que involucren el espacio muestral y los resultados obtenidos.

Palabras claves: Probabilidad, ley de los grandes números, intuiciones, experimentos aleatorios, frecuencias relativas, simulador computacional.
MSC2010: 97C10, 97K50

Effects of simulation in understanding the law of large numbers

Abstract. This paper presents some of the results in a research about the probabilistic reasoning processes of twelve students from 12 to 15 years old in a public school in a municipality in the department of Santander (Colombia) about the Law of Large Numbers. After living a direct experience, students performed simulations using *Probability Explorer*. The results suggest that to become aware of the variability in the short term and of the stability in the long one of the relative frequencies bidirectional activities involving the sample space and results must be done.

Keywords: Probability, law of large numbers, intuitions, randomized experiments, relative frequencies, computational simulator.

* Autor para correspondencia: *E-mail*: gyanez@uis.edu.co.
Recibido: 19 de febrero de 2013, Aceptado: 12 de abril de 2013.

1. Presentación

En la mayoría de los textos de matemáticas a nivel de secundaria en Colombia que tienen algún capítulo dedicado a la probabilidad [4, 24, 27], es común el uso del lenguaje formal algebraico como única forma de representación para la enseñanza de los conceptos relacionados con la probabilidad. A pesar de que los Estándares de Competencias en Matemáticas de Colombia [21] consideran los conceptos de probabilidad y estadística en la escuela desde la primaria, estos se limitan al tratamiento de los datos y, específicamente, limitan la probabilidad al solo cálculo y al empleo de algoritmos de combinatoria para resolver ejercicios. El solo uso de la probabilidad clásica no permite abordar el concepto de probabilidad en toda su extensión ni desarrollar intuiciones adecuadas [10]. La incomprensión de la probabilidad se refleja en el uso inapropiado de ciertas heurísticas (por ejemplo, las de representatividad y disponibilidad) que afectan las concepciones y juicios que emiten las personas en situaciones de incertidumbre. Lo anterior, como se ha demostrado en varias investigaciones [2, 15, 16, 17, 18, 19, 25, 26], no permite una construcción significativa de los conceptos relacionados con experimentos aleatorios; por el contrario, ha generado y reforzado algunas malas concepciones como los sesgos de equiprobabilidad, el sesgo de los valores recientes, el sesgo de desorden y de variación constante, y la insensibilidad al tamaño de la muestra, entre otros [28].

Dada la existencia de malas concepciones alrededor del concepto de probabilidad y la imposibilidad de erradicarlas con el solo enfoque clásico, la pregunta que muchos profesores se hacen es: ¿Qué metodología de trabajo es conveniente utilizar en el salón de clase para posibilitar la construcción significativa del concepto de probabilidad? Esta pregunta está, aún, lejana de responderse con precisión. Sin embargo, existen algunos elementos que pueden ayudar a acercarse a la respuesta. Según Ausubel, Novak y Hanesian [1] la construcción significativa de un concepto está ligada a la experiencia, razón por la cual se puede pensar que el trabajo experimental con fenómenos aleatorios en el salón de clase es un buen principio para la formación de estos conceptos. Ausubel *et al.* [1] proponen tener en cuenta la estructura cognitiva previa de cada estudiante para construir nuevos significados conectando de manera interactiva, y no simplemente asociativa, los conceptos preexistentes en la estructura cognitiva del sujeto con las nuevas ideas, conceptos o proposiciones. En una de las formas en que se puede generar un aprendizaje significativo tenemos el aprendizaje por descubrimiento, donde lo que va a ser aprendido no se da en su forma final, sino que debe ser reconstruido por el alumno antes de ser aprendido e incorporado significativamente en la estructura cognitiva; podemos ubicar la experiencia del estudiante como una parte fundamental en la construcción de significados.

Relacionando la anterior postura en el campo de la enseñanza de la estocástica, investigadores pioneros como Fischbein [9] han descrito las intuiciones de las personas como parte de los procesos cognitivos que intervienen en acciones mentales o prácticas del sujeto, y como mecanismos que le permiten tomar decisiones. Las personas poseen intuiciones previas o intuiciones iniciales e intuiciones secundarias. Las primeras no requieren ningún tipo de instrucción sistemática y se adquieren como parte de su experiencia. Las intuiciones secundarias se forman como consecuencia de la educación, principalmente de la escuela. Fischbein [9] sugiere que para construir significados se deben propiciar experiencias adecuadas que permitan al sujeto conectar las intuiciones primarias con las secundarias.

Para respaldar el uso de la experimentación como camino hacia aprendizajes significativos, Fischbein [9] a través de una serie de investigaciones concluyó que las personas, desde muy temprana edad, poseen una intuición primaria respecto a las frecuencias relativas asociadas a los resultados de las repeticiones de un experimento aleatorio. Más específicamente, dice:

> [. . .] para crear nuevas intuiciones correctas de probabilidad, el educando debe estar activamente involucrado en un proceso de realización de experimentos aleatorios, de adivinar resultados y evaluar posibilidades, de confrontar resultados individuales y grupales con unas predicciones realizadas a priori, etc. Nuevas intuiciones de probabilidad correctas y potentes no pueden ser producidas simplemente practicando fórmulas de probabilidad [10, p. 12].

Pero, si bien es cierto que el realizar experimentos físicos ayuda a generar una mayor comprensión alrededor del experimento aleatorio en aspectos como la identificación del espacio muestral y en la asociación ordenada entre posibilidades a priori y resultados a posteriori [28], también es cierto que dadas las pocas repeticiones que finalmente se realizan, es muy difícil que los estudiantes perciban alguna regularidad en el comportamiento de las secuencias aleatorias que permita dar algún significado a su experiencia y generar conceptos claros sobre la probabilidad de un suceso [5, 13, 23]. Una respuesta a las pocas repeticiones y a mayor cantidad de muestras la brinda la herramienta computacional.

Ahora bien, la idea de realizar experimentos es para observar la estabilidad que alrededor de su probabilidad adquieren las frecuencias relativas asociadas a cualquier evento posible cuando se aumenta el número de repeticiones, tal y como lo plantea la Ley de los Grandes Números y que se constituye en el fundamento del enfoque frecuencial de la probabilidad, considerado hoy, por una buena cantidad de educadores en probabilidad y estadística, como mucho más significativo que el enfoque clásico [12].

Un estudio que permite conocer un poco más la forma como los estudiantes razonan alrededor de esta ley es el trabajo de Pratt en [22], donde se relatan los resultados de un estudio dirigido a observar cómo un grupo de niños de diez años construían significados para la aleatoriedad en un ambiente computacional llamado Chance-Maker, cuyo entorno pedagógico particular es un dominio estocástico de abstracción. La intención del autor era estudiar las intuiciones previas de los estudiantes y analizar las posibles modificaciones o nuevos significados que los niños podían generar dentro del micromundo. Pratt estudió la conformación de *redes* formadas entre las intuiciones estocásticas de los niños con las herramientas basadas en el computador, lo que les permitió, a algunos de ellos, descubrir la relación entre la conformación del espacio muestral y las frecuencias relativas para muchas repeticiones de los experimentos aleatorios asociados.

Desde otros puntos de vista existen otras investigaciones como las realizadas en [5, 13, 14, 20, 23, 28], que muestran que el establecimiento de esa relación entre la distribución del espacio muestral y las frecuencias relativas tiene una serie de dificultades que, muchas veces, no son tan fáciles de superar. La comprensión de la Ley de los Grandes Números exige la comprensión previa de otros elementos que van más allá de la simple asociación entre la distribución de los resultados posibles y la estabilidad de las frecuencias relativas, en el sentido de que la relación que existe entre los resultados posibles se trasmite a las frecuencias relativas; esto es, si en una urna hay más bolas rojas que azules, entonces al realizar extracciones saldrán más bolas rojas que azules.

Como justificaremos más adelante con evidencia empírica, la noción de probabilidad y la Ley de los Grandes Números implican por parte del aprendiz el reconocimiento de *la variabilidad de los resultados asociados a la repetición de un experimento aleatorio, el significado de estabilidad de las frecuencias relativas y la relación entre la distribución de los resultados en el espacio muestral y el valor de probabilidad*.

Yáñez [28] plantea la necesidad de realizar experiencias que ayuden a generar comprensión alrededor del experimento aleatorio, pero dadas las pocas repeticiones que se realizan es imposible con una sola práctica obtener una buena estimación de los resultados posibles y percibir el comportamiento de las secuencias aleatorias que permitan corregir algunas malas concepciones. Para suplir esta carencia de suficientes experiencias, el autor hace uso del computador, donde es posible, con simulación, representar el experimento aleatorio y realizar muchos ensayos en un corto espacio de tiempo. Concluye en su investigación que la simulación computacional de la probabilidad permite superar algunos de los sesgos o malas concepciones que los estudiantes poseían sobre las secuencias aleatorias o sobre el valor de las probabilidades en experimentos compuestos, como el sesgo de los valores recientes, el sesgo de desorden, el sesgo de equiprobabilidad y la concepción de la variación constante de las frecuencias relativas. Pero resalta las dificultades de los estudiantes para alcanzar un nivel de comprensión suficientemente claro del concepto de probabilidad frecuencial, así como para desarrollar un concepto de la simulación computacional de probabilidad suficientemente sólido que lo hiciera independiente de los resultados en otras representaciones. También destaca la dificultad que se presenta para estimar el valor de probabilidad a partir de las frecuencias relativas, ya que muchos estudiantes optan por aceptar el último valor obtenido haciendo caso omiso de la tendencia general que presentan las frecuencias relativas al aumentar el número de repeticiones del experimento.

Por su parte, Reátiga [23] reporta en su trabajo que la experimentación física es necesaria para corregir errores en las concepciones relacionadas con el azar y la naturaleza de las pruebas experimentales, ya que permiten confrontar sus creencias con la realidad. Pero al mismo tiempo destaca que dichas experiencias no son suficientes para comprender aspectos relacionados con los espacios muestrales, las relaciones entre los resultados individuales y los patrones de las frecuencias relativas, ni para detectar la regularidad de los resultados a largo plazo, dado el fuerte pensamiento determinístico de los estudiantes.

Por estas razones surgió el interés en adoptar una herramienta computacional con un grupo de estudiantes entre los 12 y 15 años de edad, para estudiar los efectos de la simulación computacional en el aprendizaje de la Ley de los Grandes Números a partir de una situación-problema contextualizada y el uso de un simulador aleatorio llamado *Probability Explorer* (ver `http://probexplorer.com/`). Precisamente en este artículo presentamos parte de los resultados obtenidos.

Drier [6, 7, 8] justifica el uso y la necesidad de la herramienta computacional *Probability Explorer* como un micromundo, en el cual los niños usan representaciones icónicas para mostrar e interpretar datos en forma estática y como objetos dinámicos durante la simulación de experimentos aleatorios, y donde es posible que los estudiantes puedan desarrollar un pensamiento probabilístico. Este *software* fue desarrollado por la misma autora con base en una teoría constructivista de aprendizaje e investigación sobre el uso de micromundos con estudiantes. Para la autora, el ambiente computacional, las actividades significativas, instruccionales y lúdicas, junto con la motivación de los estudiantes y las interacciones sociales, operan interactivamente como agentes de construcción significativa de conceptos probabilísticos en los estudiantes. El micromundo computacional ofrece la posibilidad de desarrollar las concepciones de cada estudiante, confrontándolas en forma inmediata con los resultados de las experiencias y permitiendo confirmar o refutar sus intuiciones a través de perturbaciones mediadas por el docente y la herramienta. Este ambiente no fue creado para enseñar probabilidad a los estudiantes; en lugar de eso, el ambiente es un escenario cuasi tangible donde los estudiantes pueden desarrollar habilidades de razonamiento probabilístico a medida que utilizan las herramientas disponibles en el micromundo para diseñar sus propios experimentos y analizar datos generados aleatoriamente. La intención de Drier y la nuestra no era reemplazar las experiencias físicas con simulaciones digitales, sino establecer conexiones significativas entre los dispositivos físicos y los dispositivos virtuales para que dicha transición fuera lo más natural posible.

Drier resalta la ventaja de amplificación del instrumento de mediación, que permite a los estudiantes analizar los resultados experimentales de forma visual al comparar las cantidades y observando el orden en que ocurren los resultados, desarrollando habilidades en la comprensión de múltiples representaciones que cambian instantáneamente y con propiedades que los dispositivos físicos o computacionales difícilmente pueden ofrecer. En este caso, el micromundo del computador no sólo se conecta con el mundo físico del niño a través del lenguaje icónico dinámico, sino que extiende sus acciones potencialmente disponibles en este nuevo "mundo" matemático, donde el uso de estas representaciones como unidades dinámicas de análisis permiten al niño visualizar conceptos como la Ley de los Grandes Números y desarrollar usos apropiados de la heurística de representatividad.

Las representaciones disponibles en el micromundo *Probability Explorer* son representaciones multi-enlazadas, que se actualizan simultáneamente cuando los eventos aleatorios son simulados. Estas representaciones incluyen representaciones de íconos movibles, arrastrando con el cursor. Además, en él se puede manipular el número de pruebas (de 1 en 1 ó 20 pruebas seguidas, por ejemplo), las cuales se ejecutan con solo un clic en el icono "correr el experimento" (*Run Experiment*), representado por un muñeco que corre. El usuario también puede manipular la velocidad de la simulación en el menú de opciones (*Simulation Speed*) en tres niveles: baja (*slow*), media (*medium*) y rápida (*fast*). Incluye también íconos para crear su propio experimento con dibujos alusivos (balones de diferentes deportes, figuras de estado del tiempo, población, caras, caras de dado de 1 a 9).

En la Figura 1 se señalan en español varios de los íconos; se incluye un block de notas (*notebook*) que permite al estudiante describir las conclusiones a que puede llegar en uno o varios experimentos y pegar tablas de resultados o gráficos. El programa permite guardar cada sesión del estudiante como un archivo, con el propósito de poder analizar los experimentos realizados por cada uno de ellos. En la Figura 2 se ilustran las diferentes opciones que tiene el programa para diseñar un experimento aleatorio (opciones del clima, dados de hasta nueve caras, deportes, población, figuras geométricas, etc.). En la Figura 3 se ilustra un ejemplo de una simulación del lanzamiento de dos monedas "justas" con 200 repeticiones, donde se puede ver la lista de resultados para analizar la secuencia aleatoria, la tabla de frecuencias absolutas y relativas (en forma decimal y porcentual), así como la representación gráfica de resultados en un diagrama de barras.

En el apartado siguiente se presenta brevemente la metodología de la investigación implementada. Posteriormente se realiza el análisis de los resultados, donde, en particular,

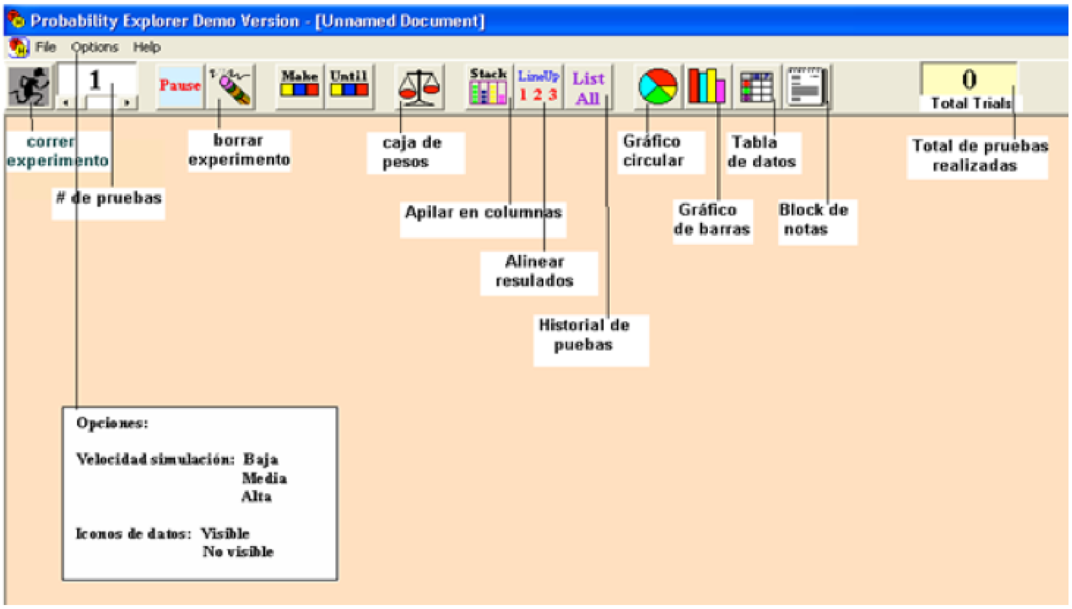


Figura 1. Iconos de *Probability Explorer.*

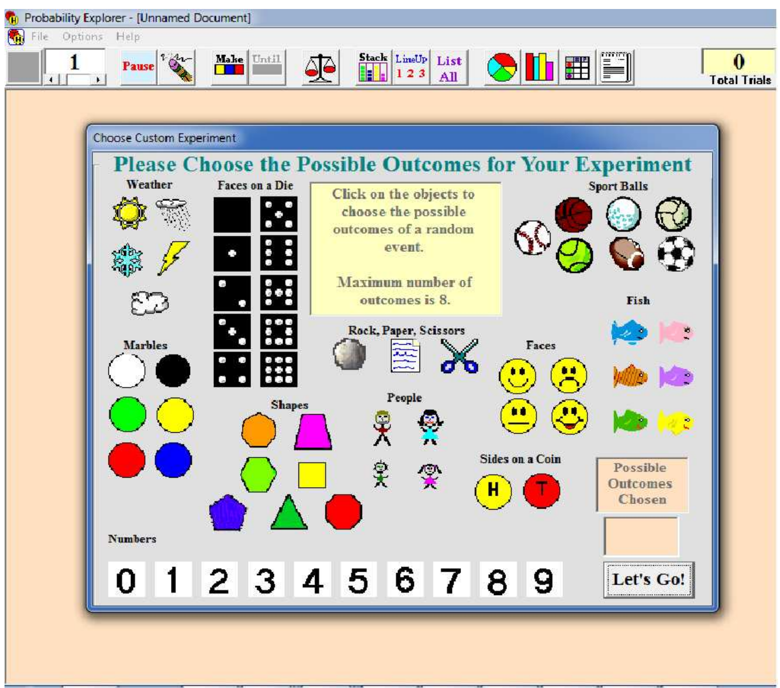


Figura 2. Opciones de experimento.

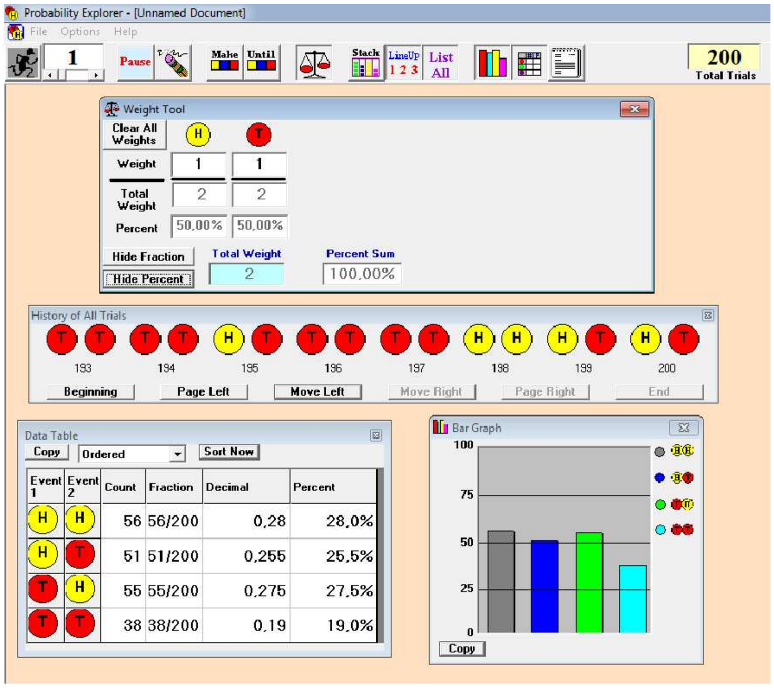


Figura 3. Ejemplos de simulación.

se presentan algunos apartes de la conversación que se tuvo con Camilo (uno de los estudiantes) mientras daba respuesta al problema que se le propuso y que debía resolver con la ayuda del computador en la fase de evaluación al final de la experiencia. Finalmente, se presentan las conclusiones de este trabajo relacionado con el proceso de formación de significados asociados a la Ley de los Grandes Números experimentados por uno de los estudiantes que participaron en esta experiencia.

2. Metodología

La investigación realizada es cualitativa, ya que más que cifras intentamos explicar los procesos mentales vividos por los estudiantes a través de las actividades realizadas. En particular, este estudio se limita a referir los procesos vividos por un estudiante que se adoptó como caso. La muestra objeto de estudio fue un grupo de 40 estudiantes de octavo grado entre los 12 y 15 años del Instituto Técnico Industrial de Puente Nacional (zona rural en el departamento de Santander, Colombia), los cuales no habían recibido ningún tipo de enseñanza formal en los conceptos de probabilidad y que tampoco fueron sometidos a ningún tipo de instrucción en el desarrollo de la investigación. La duración del proyecto fue de un año, y su etapa de aplicación fue de dos meses aproximadamente (8 sesiones, 24 horas en total). La recolección de información se hizo a través de los talleres escritos de cada actividad, los archivos de computador y entrevistas individuales videograbadas de cada actividad (20 horas en total). Finalmente se editó un vídeo de la experiencia, que resumió el proceso de investigación en veinte minutos y algunos de los resultados encontrados. Durante el trabajo de campo con los estudiantes se desarrolló una serie de actividades en doce momentos, que comenzaron con un análisis diagnóstico de las concepciones de los estudiantes respecto a las secuencias aleatorias, y que continuaron a partir de los resultados experimentales obtenidos de una promoción de helados que se realizó en el colegio, la cual giraba en torno a una apuesta relacionada con el resultado del lanzamiento de una moneda: si el cliente acertaba en el resultado, se ganaba un helado; si no acertaba, debía pagar el valor del helado sin recibirlo. Inicialmente se dio a conocer la promoción a los estudiantes para que analizaran su conveniencia a largo plazo. Posteriormente, los mismos estudiantes efectuaron la promoción en la hora de recreo. Con los resultados obtenidos, se les pidió a los estudiantes que analizaran lo ocurrido y lo confrontaran con sus análisis previos. La necesidad de mayores datos condujo a los estudiantes a diseñar modelos que representaran la situación real y que permitieran su simulación tanto física como computacional. El modelo físico consistió en reemplazar la escogencia del cliente por el lanzamiento de otra moneda, de tal forma que si los resultados de las dos monedas coincidían el concursante ganaba el helado, en caso contrario, debía pagarlo.

Después de experimentar repetidas veces, tanto física como computacionalmente, los estudiantes diseñaron una nueva promoción que garantizaba ganancias utilizando otro dispositivo aleatorio: la extracción de balotas de una bolsa que contenía bolas de dos colores, 2 bolas rojas y 3 bolas amarillas, donde el cliente ganaba si extraía la bola roja y perdía si extraía la amarilla. Esta nueva situación aleatoria, no equiprobable, también fue desarrollada a través del análisis y recolección de datos en una experimentación física y, luego, confrontada con los resultados de la simulación.

Todas las actividades se centraban en el análisis de los resultados que, si bien en un principio realizaron solamente a partir de los datos numéricos, luego pudieron hacerlo a través de las múltiples representaciones dinámicas ofrecidas por el micromundo *Probability Explorer*. La confrontación de respuestas y sus discusiones, en los diversos momentos de socialización entre estudiantes y docentes investigadores, permitieron el esclarecimiento de los significados que los estudiantes se fueron formando a través de las actividades desarrolladas. Por último, se realizó una evaluación escrita y una computacional para contrastar sus respuestas con el diagnóstico inicial y poner a prueba las nuevas intuiciones y significados alrededor de la Ley de los Grandes Números. La evaluación computacional consistió de un solo problema: estimar el número de bolas de dos colores distintos que contiene una urna utilizando solo extracciones con sustitución. En la evaluación escrita se pidió a los estudiantes que interpretaran los valores de probabilidad de diversos fenómenos aleatorios en términos de los posibles resultados para pocas y muchas repeticiones. En la prueba diagnóstica y en la actividad de experimentación real se trabajó con toda la muestra, pero después de analizar los resultados del diagnóstico, el interés de los estudiantes y la disposición de los padres de familia para permitir a los estudiantes asistir a la institución en horario extra clase decayeron, por lo cual el grupo se redujo a doce estudiantes que desarrollaron la totalidad de las actividades propuestas, exceptuando la fase final de evaluación, que solo se efectuó con dos estudiantes. Si bien se tienen resultados que evidencian los cambios sufridos por los estudiantes en sus concepciones alrededor de la probabilidad, sólo presentamos aquí los resultados que nos permiten establecer si la estrategia aplicada fue útil a la hora de construir el sentido de la Ley de los Grandes Números y formar un significado adecuado del concepto de probabilidad. En especial, tuvimos en cuenta el proceso con el cual los estudiantes construyeron su red de significados a partir de intuiciones relacionadas con las frecuencias relativas mediante la coordinación de significados.

La naturaleza de los problemas planteados a los estudiantes está relacionada con el origen de la Ley de los Grandes Números en la obra de Bernoulli: estimar la probabilidad cuando se conocen los resultados.

> La progresiva estabilización de las frecuencias relativas de un resultado dado en un número grande de experimentos, que ha sido observada durante siglos y fue expresada por Bernoulli como un teorema matemático, sirvió como justificación para la definición frecuencial de la probabilidad. [...] Esta idea no está libre de dificultades, debido a que la naturaleza específica de la convergencia aleatoria es difícil de comprender, y la aparición de largas rachas, coincidencias y patrones inesperados son contra-intuitivos [3, p. 30].

De modo que el estudiante debe analizar el comportamiento de las frecuencias relativas a medida que aumenta el número de lanzamientos o extracciones hasta encontrar el valor de probabilidad por aproximación. Y como mencionan Batanero *et al.* [3], estas frecuencias pueden converger lentamente y mostrar fluctuaciones, aspecto que muchos estudiantes evitan recurriendo a la estrategia del último valor generado [28]. El uso crítico de una herramienta computacional que permita generar muchos experimentos aleatorios en gran número de pruebas puede ayudar a encontrar dicha aproximación y relacionarla con la probabilidad teórica y el espacio muestral para generar un concepto significativo.

3. Análisis de resultados

Dada la extensión y riqueza de los resultados obtenidos en esta investigación, y recordando que solo se presentarán los resultados obtenidos respecto a la construcción del significado conceptual de la Ley de los Grandes Números, solo se transcriben algunos apartes de la conversación realizada con Camilo (uno de los estudiantes seleccionados como caso) cuando trabajaba en el computador e intentaba dar respuesta a la pregunta de conocer el número exacto de bolas que estaban dentro de una urna que contenía seis bolas negras y dos bolas azules.

Camilo comenzó simulando pocas extracciones, una por una, con la esperanza de encontrar el número de balotas de la urna en pocas extracciones. Cuando había realizado 14 extracciones observó los resultados y los porcentajes para cada color: Cinco azules (35,71 %) y nueve negras (64,29 %). Luego borró los resultados y simuló el experimento ejecutando directamente 10 extracciones y obtuvo dos azules (20 %) y ocho negras (80 %).

C (Camilo): Pues será una bola azul y cuatro negras.

Camilo cae en el sesgo conocido como la *ley de los pequeños números* [16], olvidando momentáneamente el significado de variabilidad a corto plazo del cual había dado muestras de reconocer en las experimentaciones realizadas previamente. Este olvido es una evidencia de que el significado creado de "variabilidad" en los resultados a corto plazo, aún está conectado débilmente con el significado de "estabilidad" de los resultados a largo plazo.

...

P (Profesor): ¿Y si vuelve a hacer otras 10 extracciones?

C: 5 azules (50 %) y 5 negras (50 %).

P: ¿Por qué escogió el 10?

C: Pues porque no creo que la bolsa tenga más de 10 balotas.

P: Entonces, ¿el número de extracciones que debe realizar depende del número de balotas?

C: Pues no, solamente que con 10 hay más posibilidades de que me caigan exactos los porcentajes, porque si hago por ejemplo 100 extracciones me puede salir 50,3 % de negras o azules; entonces eso no me conviene, porque necesito saber el número exacto de balotas.

Si bien con 100 no era el caso, porque los porcentajes coinciden con el resultado, Camilo intuía que las proporciones exactas le permitirían acercarse mejor a la composición de la urna, ya que el número de bolas es un número entero. Por eso, asume 10 extracciones que, aunque después lo niegue, cree que le deben reportar valores más cercanos, pensando tal vez en una relación cuasibiyectiva entre bolas y resultados a obtener. Este aspecto no fue percibido por el profesor, que, acto seguido, le sugiere que realice 100 extracciones.

P: Muestre con 100 extracciones.

C: Se ve que salen porcentajes con decimales [lo dice porque el programa va mostrando los respectivos porcentajes a medida que se va aumentando el número de extracciones], y necesito dar un número exacto de balotas [refiriéndose a los valores que toman las frecuencias a medida que aumenta el número de extracciones].

Al finalizar la simulación con 100 extracciones se obtuvo 29 bolas azules (29 %) y 71 bolas negras (71 %), que se reflejaron en las representaciones tabular y gráfica de la pantalla.

P: ¿Qué sucedió ahí con las 100 extracciones?

C: Parece que hubiera 3 bolas negras y 1 bola azul, ó 6 negras y 2 azules.

...

Borró los resultados y volvió a simular con 100 extracciones.

C: Mire, 73 % de bolas negras y 27 % azules. O sea, si hubiera cuatro balotas en la bolsa, entonces 3 bolas serían negras y 1 azul.

Nuevamente simuló otras 100 extracciones.

C: Mire, 24 % para las azules y 76 % para las negras ($A = 24$ y $N = 76$).

P: ¿Como podría estar seguro de esa relación que usted dijo (1 bola azul y 3 negras)?

C: No estoy seguro completamente... ¿únicamente simulando?

Aquí el profesor decide dar una ayuda a Camilo con el ánimo de recordarle significados adquiridos previamente.

> *P:* Sí. Por ejemplo: abra otra ventana y configure una urna de la cual usted sepa cómo va a salir el gráfico.

Camilo configuró en otra ventana una urna con 1 bola roja y 2 amarillas, y ejecutó 100 extracciones.

> *P:* ¿Cómo cree que va a salir el gráfico?
>
> *C:* Una tercera parte para la roja y dos terceras partes para la amarilla.

Ejecutó la simulación y abrió solo el gráfico circular para su análisis.

> ...
>
> *C:* Entonces yo siempre he creído que el programa trata como de basarse en eso, entonces de pronto... no pero... ¿para hacer más extracciones?

Borró los resultados anteriores y ejecutó 250 extracciones con velocidad lenta, observando detalladamente los cambios en las múltiples representaciones.

> *P:* ¿Qué está sucediendo ahí? [Mientras el simulador genera resultados con baja velocidad].
>
> *C:* Pues que la balota roja intenta irse mucho para acá (ser menor de $^1/_4$ de circunferencia), algo que no debería pasar. Pero me estoy dando cuenta de que con más lanzamientos trata de estabilizarse más en los valores que yo dije, es decir $^1/_3$ de las rojas. ¿Sí ve que al principio nos daba valores que no correspondían? En cambio con muchas extracciones trata de dar más preciso, donde debe ser.

Al realizar el proceso contrario: conocida la composición de la urna describir el comportamiento de las extracciones, Camilo logra coordinar los significados que relacionan el número de extracciones y el valor de probabilidad.

> *P:* Entonces, ¿qué cree que debe hacer para averiguar el número de balotas negras y azules de la urna?
>
> *C:* Hacer la simulación por ahí con 250 extracciones, como aquí.
>
> *P:* Esperemos a ver qué resultados obtiene en esta ventana. ¿El gráfico qué refleja?
>
> *C:* Hay como una tercera parte de bolas rojas, pero un poco corrida (inexacta).
>
> *P:* ¿Qué pasa a medida que aumentan las extracciones?
>
> *C:* Trata de irse más para donde debe, o sea a $^1/_3$; me dio 29,2 % [refiriéndose a la frecuencia de balotas rojas]. Para que de más o menos exacto no sé cómo hacer.
>
> *P:* ¿Está mejor que con las 50 extracciones, o no?

C: Pues un poco, sí, porque con las 50 se pasó un 5 % y aquí ya un 4 % no fue mucho la diferencia pero mejoró un poco [refiriéndose la frecuencia de balotas rojas].

P: Entonces, ¿qué más puede hacer?

C: No sé, porque seguir aumentando la cantidad de extracciones no es que me haya dado resultado.

P: ¿Cree que 250 extracciones son suficientes para que le dé lo esperado?

C: Sería probar de otra forma.

...

P: ¿Para cuántas extracciones?

C: 366 extracciones.

P: ¿Qué pasa si continúa la simulación con la urna?

C: Vemos que está como cerca, o sea como que con muchas extracciones es más exacto. Digamos que está la urna en la realidad y vamos a hacer 5 extracciones; entonces sacamos una por una las balotas y puede que en estas cinco primeras extracciones salga la misma bola y luego empiecen a salir las otras. A largo plazo nos vamos dando cuenta de los porcentajes, porque al principio sí probábamos con pocas extracciones; no va a dar como queríamos, pero con muchas extracciones, aquí llevamos 520 [refiriéndose a la simulación que continuó ejecutando], ya los porcentajes nos van a convenir más.

P: ¿Nos están conviniendo?

C: Es decir, que de acuerdo a mi urna el resultado de balotas rojas ya es de $^1/_3$.

P: ¿Cuál es la probabilidad de sacar una bola roja?

C: Un tercio.

P: O sea, que si voy hacer 10 extracciones, ¿cuántas bolas cree que van a salir rojas?

C: De 10 extracciones por ahí 3 o 4 bolas rojas. Es que con poquitas extracciones ya es más difícil, porque uno va a dar el porcentaje.

P: ¿Cuándo funciona el valor de la probabilidad?

C: Cuando es una gran cantidad.

Detallemos que el estudiante aún no logra coordinar por completo el significado de "variabilidad" y el de "probabilidad". Es decir, no logra explicar el por qué no se refleja el valor de probabilidad en los resultados aleatorios a corto plazo. Es decir, complementa el significado global anterior así: "Si el valor de probabilidad no se refleja a corto plazo en las frecuencias relativas, entonces el valor de probabilidad falla a corto plazo". Sin embargo, logra coordinar correctamente el significado de "estabilidad" y "probabilidad", al explicar cómo se genera esa estabilidad a partir de los cambios que se presentan en las secuencias aleatorias a medida que se incrementan las pruebas.

...

P: Volviendo a nuestro problema.

C: Pues aquí con 750 extracciones nos dimos cuenta que fue más exacto. Obtuvo en la tabla de frecuencias las siguientes cantidades de la urna configurada por él con una bola roja y dos amarillas: rojas el 33,2 % y amarillas el 66,8 %.

C: Entonces sería volver a la otra ventana y probar con 750 extracciones.

C: Acá nos dio 24,67 % de bolas azules, que aproximando nos da 25 %; y 75,33 % de bolas negras, que aproximando daría 75 %. Entonces yo digo que el 25 % de las balotas que hay en bolsa son azules y 75 % son negras.

P: Entonces, ¿cuántas bolas podrían haber en la urna?

C: Pues podría haber 1 y 3 ó 2 y 6 y así sucesivamente. Por cada bola azul hay tres negras. ¿Si?

P: ¿Podría asegurar eso?

C: Sí, porque con este experimento nos dimos cuenta de que casi funciona; o sea, cuanto más lanzamientos, más exacta será la respuesta.

P: ¿Qué pasaría si se duplican a 1.500 extracciones?

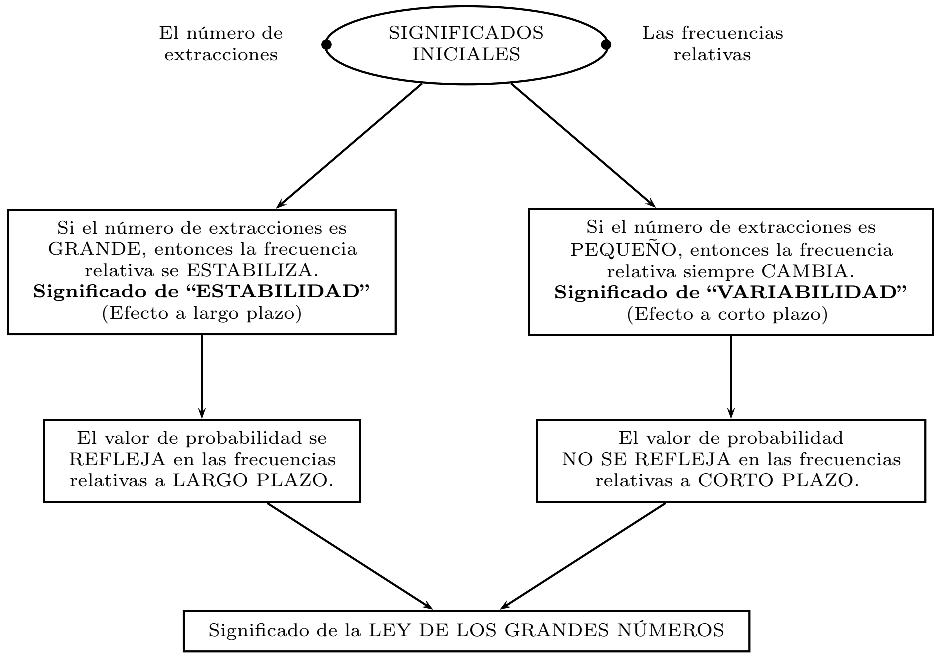
C: Pues sería aun más exacta. Ahí ya hay posibilidades de que saliera 25 y 75 %, o con una diferencia mínima.

P: Entonces, ¿cuándo funciona la probabilidad de un evento?

C: Pues mientras más largo el tiempo o lo que se esté manejando, mientras mayor sea la cantidad, pues más exacta es.

El proceso inverso de estimar la distribución de bolas en la urna, que para el caso no es otra cosa que el valor del parámetro p (la probabilidad de éxito) en un modelo binomial, nos permitió detectar debilidades en las conexiones establecidas por Camilo en las actividades previas, y darnos cuenta, cómo a través de este proceso, el mismo estudiante, con el acompañamiento del docente, quien lo conduce a la situación donde se conoce el contenido de la urna para que observe lo que sucede con la distribución de las extracciones realizadas, afianza y coordina eficazmente los significados para generar una comprensión de la Ley de los Grandes Números, que finalmente le permitió resolver la situación. Ahora sí podemos evidenciar en el estudiante una comprensión completa del significado de la Ley de los Grandes Números a partir de la coordinación de los significados de "variabilidad" y "estabilidad" de los resultados de un experimento aleatorio a corto y largo plazo.

El siguiente esquema muestra la forma en que se generaron y coordinaron los significados a través de la evaluación computacional. Así, Camilo descubre la verdadera composición de la urna: dos bolas azules y seis negras; generalizando el resultado, por cada bola azul hay tres negras, lo cual reflejan un nivel más elevado de comprensión de la solución del problema.



Vale la pena resaltar el hecho de que Camilo, si bien intentó inicialmente responder con valores enteros para la cantidad de bolas de cada color en la urna, después de realizar repetidas simulaciones cambia su estrategia y empieza a trabajar con las frecuencias relativas, aspecto este que le permite darse cuenta de que el problema en sí mismo no tiene solución única; lo único que se puede deducir es la razón entre la cantidad de bolas de los colores presentes ("Pues podría haber 1 y 3 ó 2 y 6, y así sucesivamente. Por cada bola azul hay tres negras"), o, lo que es equivalente, la razón de las bolas de cada color respecto al total de bolas, es decir, el valor de probabilidad.

En resumen, el planteamiento del problema que indagaba por la composición de la urna conociendo las frecuencias relativas, contrario al que se había realizado previamente, como era predecir el comportamiento de las frecuencias relativas conociendo la composición de la urna, nos permitió ver que las conexiones entre los significados generados por el estudiante no eran lo suficientemente fuertes y presentaban inconsistencias (como la ausencia del significado de "variabilidad" a corto plazo). Fue precisamente el volver al problema directo lo que tal vez catapultó en Camilo la comprensión del significado total de la ley de los grandes números permitiéndole estimar el contenido de la urna.

4. Conclusiones

La construcción conceptual de la Ley de los Grandes Números y el concepto de probabilidad asociado al enfoque frecuencial de la probabilidad requieren, a nuestro juicio, de tres significados básicos, como son: la variabilidad de los resultados obtenidos cuando se repite un experimento aleatorio; la estabilidad de las frecuencias relativas asociadas a los resultados de un evento, y la relación entre el valor límite de esas frecuencias con la distribución de los resultados posibles en el espacio muestral y el valor de probabilidad. A continuación analizamos cada uno de estos significados.

- *El surgimiento del significado de variabilidad.* Si bien percibir la variabilidad de los resultados de un experimento aleatorio no tiene mayor complicación, sí es un problema cuando solamente se posee este significado, ya que puede dar lugar a pensar que esta es sencillamente el reflejo del azar en el sentido de su impredecibilidad y de la imposibilidad de control. Muchas de las personas que asumen la variabilidad en este sentido pueden terminar pensando que si nada se puede predecir, entonces cualquier resultado es igualmente posible, cayendo así en el sesgo de equiprobabilidad. La experiencia realizada con los estudiantes mostró que estos solo entendieron la variabilidad cuando la confrontaron con el concepto de estabilidad. En pocas palabras, para que el estudiante pudiera construir el significado claro de "variabilidad" era necesario primero el surgimiento del significado de "estabilidad". Es el juego dialéctico entre una propiedad y su antagónica lo que permite entenderlas. Es semejante a la idea genética de Piaget respecto a la aleatoriedad, cuando afirmaba que para que el muchacho comprenda la irreversibilidad de los fenómenos aleatorios y su carácter de impredecibilidad, es absolutamente necesario que comprenda totalmente la relación causal. Es el manejo de la contradicción: se requiere lo contrario para comprender la propiedad inicial; se debe contar con elementos que no cumplan la propiedad para percibir la propiedad en sí misma; no basta que existan objetos que la posean: es necesario que existan los que no la poseen.

- *El significado de estabilidad.* Este significado requiere tener en cuenta dos aspectos para su formación: tamaño y repeticiones. Tamaño hace referencia al número de pruebas que se realicen, es decir, al tamaño muestral indicado y que se toma como base del análisis; repeticiones se refiere tanto a la realización de más pruebas como a la repetición de las muestras. Se trata, en últimas, de la variación horizontal y vertical que describe Yáñez [28]: la horizontal se refiere al cambio de los valores de las frecuencias relativas a medida que se aumenta el número de repeticiones; la vertical se refiere a los cambios que sufren los valores de las frecuencias relativas cuando se cambia la muestra permaneciendo constante el número de repeticiones y que son modelados con la distribución binomial. Con la horizontal se visualiza el acercamiento a cierto valor, y con la vertical se corrobora que este acercamiento es independiente de la muestra tomada.

  Las evidencias obtenidas reflejan la comprensión por parte de los estudiantes de la componente horizontal de la estabilidad. Este significado surgió de manera más efectiva en los estudiantes que estuvieron bajo el continuo seguimiento a través de las entrevistas; es decir, al parecer la necesidad de expresar verbalmente las respuestas a las preguntas al interactuar con el docente investigador, de alguna manera permite al estudiante "ver" y expresar lo que antes no era posible cuando las expresaba por escrito.

- *Relación del espacio muestral y del valor de probabilidad con las frecuencias relativas de los resultados obtenidos.* En el caso de uno de los estudiantes evaluados al final (Camilo), este significado surge de manera significativa cuando analiza las condiciones de la nueva promoción con un espacio no equiprobable, y de alguna manera predice que la probabilidad de sacar una bola roja depende de la cantidad de bolas que hay en la bolsa: "Si hay dos bolas rojas y 3 amarillas, entonces la probabilidad de ganar es del 40 %, y de perder de 60 %". Esta conjetura fue rápidamente validada por los resultados de la experimentación física y respaldada por la simulación computacional, lo que le permitió asociar a cada espacio muestral dos valores de probabilidad: los correspondientes a cada color de bola. Con este nuevo significado fue más sencillo para Camilo dar sentido al significado de estabilidad de las frecuencias relativas, centrando su atención en buscar la forma de encontrar el valor de probabilidad, en los resultados de las simulaciones. Para ello, Camilo relacionó el espacio muestral con el valor de probabilidad y a su vez relacionó el valor de probabilidad con la frecuencia relativa al condicionar el "cumplimiento" del valor de probabilidad al hecho de que dicho valor debía reflejarse en las frecuencias porcentuales, encontrando que esta condición sólo se cumplía al realizar el experimento a largo plazo, como lo pudo constatar en el simulador. Esta evidencia de comprensión de la Ley de los Grandes Números por parte de Camilo fue ratificada en la prueba final escrita, cuando se le presentaron espacios muestrales finitos y se le pidió que realizara sus predicciones a corto, mediano y largo plazo.

Referencias

[1] Ausubel D.P., Novak J.D. and Hanesian H., *Educational Psychology: A Cognitive View, 2 ed.,* Warbel & Peck, New York, 1986.

[2] Batanero C. and Sánchez E., "What is the nature of high school student's conceptions and misconceptions about probability?," in *Exploring probability in school: challenges for teaching and learning* (ed. Jones G.), Kluwer Academic Publisher, (2005) 260–289.

[3] Batanero C., Henry M. and Parzysz B., "The nature of chance and probability", in *Exploring probability in school. Challenge for teaching and learning* (ed. Jones J.A.), Springer, New York, (2005) 15–37.

[4] Bermúdez M. *Matemática activa Pitágoras 7*, Ed. Santillana, Bogotá, 2005.

[5] Bohórquez J.S. y Zárate J.M., "Abstracciones situadas en un entorno experimental para la comprensión de la ley de los grandes números en niños de quinto de primaria", Thesis (Trabajo de Grado Licenciatura en Matemáticas), Universidad Industrial de Santander, Bucaramanga, 2009.

[6] Drier H.S., "Children's Probabilistic Reasoning with a Computer Microworld", Thesis (Ph.D.), University of Virginia, 2000.

[7] Drier H.S., "The Probability Explorer: A research-based microworld to enhance children's intuitive understandings of chance and data", *Focus on Learning Problems in Mathematics* 22 (2000), no. 1-2, 165–178.

[8] Drier H.S., "Children's meaning-making activity with dynamic multiple representations in a probability microworld", in *Proceedings of the twenty-second annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (ed. Fernández M.), Tucson, AZ (2000), 691–696.

[9] Fischbein E., *The Intuitive sources of Probability Thinking in Children*, Dordrecht: Reidel, 1975.

[10] Fischbein E., "Intuition and proof", *For the Learning of Mathematics* 3 (1982), no. 2, 9–19.

[11] Godino J.D. y Batanero C., "Significado institucional y personal de los objetos matemáticos", *Recherches en Didactiques des Mathématiques* 14 (1994), no. 3, 325–355.

[12] Godino J., Roa R., Recio A., Ruiz F. y Pareja J., "Análisis didáctico de un proceso de estudio de la ley empírica de los grandes números", in *Investigaciones Actuales en Educación Estadística y Formación de Profesores* (ed. Ortiz J.), Departamento de Didáctica de la Matemática, Granada (2011), 9–22.

[13] Jaimes E. y Martínez J., *Probability Explorer: Un socio cognitivo en la construcción del significado de la Ley de los Grandes Números con estudiantes de octavo grado en el Instituto Técnico Industrial de Puente Nacional*, Thesis (Tesis de de especialización en Educación Matemática), Universidad Industrial de Santander, Bucaramanga, Colombia, 2007.

[14] Jaimes E., *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 en Ciencias con Especialidad en Matemática Educativa, no publicada, Centro de Investigaciones y Estudios Avanzados del Instituto Politécnico Nacional, Cinvestav-IPN, México, D.F., 2011.

[15] Kahneman D. and Tversky A., "Subjective probability: A judgment of representativeness", *Cognitive Psychology* 3 (1972), 430–454.

[16] Kahneman D., Slovic P. and Tversky A., *Reasoning under uncertainty: Heuristic and biases*, Cambridge University, 1982.

[17] Konold C., "Informal conceptions of probability", *Cognition and Instruction* 6 (1989), 59–98.

[18] Konold C., "Understanding students' beliefs about probability", in *Radical Constructivism in Mathematics Education* (ed. von Glasersfeld E.), Dordrecht: Kluwer (1991), 139–156.

[19] Lecoutre M.P., "Cognitive Models and problem spaces in "purely random situations", *Educational Studies in Mathematics* 23 (1992), 557–568.

[20] Mantilla M. y Martínez M., "Construcción de significados del concepto de probabilidad frecuencial en un ambiente computacional. Una experiencia con profesores en formación", Thesis (Trabajo de Grado Licenciatura en Matemáticas), Universidad Industrial de Santander, Bucaramanga, 2007.

[21] MEN, *Estándares Básicos de Competencias en Lenguaje, Matemáticas, Ciencias y Ciudadanas*, Ministerio de Educación Nacional, 46–95. Bogotá, Colombia, 2006.

[22] Pratt D., "The Co-ordination of Meanings for Randomness", *For the Learning of Mathematics* 18 (1998), no. 3, 2–11.

[23] Reátiga A., "Confrontación entre realidad y modelo teórico: Una propuesta para desarrollar la intuición probabilística en los niños de sexto grado", Thesis (Tesis de especialización), Universidad Industrial de Santander, Bucaramanga, 2004.

[24] Rodríguez B., Dimaté M. y Beltrán L., *Matemáticas*, Prentice Hall 9, Pearson Educación de Colombia, Bogotá, 2000.

[25] Serrano L., Batanero C., Ortiz J.J. y Cañizares M.J., "Heurísticas y sesgos en el razonamiento probabilístico de los estudiantes de secundaria", *Educación Matemática* 10 (1998), no. 1, 7–26.

[26] Tversky A. and Kahneman D., "Availability: A Heuristic for Judging Frequency and Probability", *Cognitive Psychology* 5 (1973), 207–232.

[27] Vega G. y Cely J., *Fórmula 10°: Trigonometría y Geometría Analítica*, Ed. Voluntad, Bogotá, 2009.

[28] Yáñez G., *Estudios sobre el papel de la simulación computacional en la comprensión de las secuencias aleatorias, la probabilidad y la probabilidad condicional*, Theis (Ph.D.) Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional Cinvestav-IPN, México D.F., 2003.