Educación

Sección a cargo de

María Luz Callejo

El siguiente artículo apareció publicado en el número de junio/julio de 2000 (volumen 47, número 6) de *Notices of the American Mathematical Society*. La Gaceta de la Real Sociedad Matemática Española agradece a la American Mathematical Society el permiso para la presente traducción y publicación. Lo hacemos acompañar por unos comentarios de su traductor, Juan D. Godino.

Objetivos y métodos de la investigación en Educación Matemática

por

Alan H. Schoenfeld

Bertrand Russell definió la Matemática como la ciencia en la que nunca sabemos sobre lo que estamos hablando ni si lo que decimos es verdadero. Se ha comprobado que la Matemática es ampliamente aplicable en muchos otros campos científicos. Por tanto, la mayoría de los científicos no saben de lo que están hablando ni si lo que dicen es verdad.
- Joel Cohen, *"Sobre la naturaleza de las demostraciones matemáticas"*

No hay demostraciones en Educación Matemática.
- Henry Pollak

La primera de las citas anteriores es humorística; la segunda es seria. Ambas, sin embargo, sirven para resaltar alguna de las principales diferencias entre la Matemática y la Educación Matemática –diferencias que se deben entender si se desea comprender la naturaleza de los métodos y resultados en la Educación Matemática.

La cita de Cohen señala algunos aspectos serios de la Matemática. Al describir distintas geometrías, por ejemplo, partimos de términos no definidos. A continuación, siguiendo las reglas de la lógica, probamos que si ciertas afirmaciones son verdaderas, se deben deducir otros resultados. Por otra parte, los términos son indefinidos; esto es lo que se quiere decir con que "no sabemos nunca sobre lo que estamos hablando". Sin embargo, los resultados son definitivos. Como Gertrude Stein podría haber dicho, una demostración es una demostración es una demostración.

Otras disciplinas trabajan de forma diferente. La afirmación de Pollak no supone un rechazo hacia la Educación Matemática, sino una indicación de que la naturaleza de la evidencia y la argumentación en la Educación Matemática es bastante diferente de la naturaleza de la evidencia y la argumentación en Matemáticas.

Lo cierto es que el tipo de preguntas que se pueden plantear (y a las que se pueda dar respuesta) en investigación educativa no es la que los matemáticos esperarían. Por el contrario, los matemáticos y los investigadores en Educación suelen mantener perspectivas diferentes sobre los objetivos y los fines de la investigación en Educación Matemática.

Este artículo comienza con un intento de establecer algunas de las perspectivas relevantes y proporcionar el fundamento de la naturaleza de la investigación en Educación Matemática. Entre las cuestiones exploradas se encuentran las siguientes: ¿Cuál es exactamente la finalidad? Esto es, ¿cuáles son los propósitos de la investigación en Educación Matemática? ¿Qué características tienen las teorías y los modelos en Educación comparados con los correspondientes a la Matemática y las Ciencias Físicas? ¿Qué tipo de cuestiones puede responder la investigación educativa? ¿Qué puede ser una respuesta razonable a tales cuestiones? ¿Qué tipos de evidencias son apropiadas para fundamentar las afirmaciones educativas? ¿Qué tipo de métodos pueden generar tal evidencia? ¿Qué criterios podemos tener para juzgar las afirmaciones, modelos y teorías? Como veremos, existen diferencias significativas entre la Matemática y la Educación con respecto a todas estas cuestiones.

Objetivos

La investigación en Educación Matemática tiene dos objetivos principales, uno puro y otro aplicado:

- Puro (Ciencia Básica): Comprender la naturaleza del pensamiento matemático, la enseñanza y el aprendizaje;
- Aplicado (Ingeniería): Usar esa comprensión para mejorar la instrucción en Matemáticas.

Ambos objetivos están íntimamente relacionados, siendo el primero al menos tan importante como el segundo. La razón es simple: sin una profunda comprensión del pensamiento, la enseñanza y el aprendizaje, no es posible ningún progreso consistente en el “frente aplicado”. Una analogía útil es la relación entre la investigación en Medicina y la práctica médica. Hay una gran cantidad de investigación médica. Alguna se hace de manera urgente, con aplicaciones potenciales en un futuro inmediato. Otras investigaciones se hacen con el fin de comprender los mecanismos fisiológicos básicos. A largo plazo los dos tipos de trabajo viven en sinergia. Esto es así porque el conocimiento básico es de interés intrínseco y porque establece y refuerza los fundamentos sobre los que se basa el trabajo aplicado.

Es necesario comprender estos dos propósitos duales, pues contrastan fuertemente con el propósito simple de la investigación en Educación Matemática, tal y como es vista desde la perspectiva de muchos matemáticos:

"Dígame lo que funciona en la clase".

Decir esto no implica que los matemáticos no se interesen en algún nivel abstracto por la investigación básica en la Educación Matemática, sino que su expectativa primaria es la utilidad en términos más bien directos y prácticos. Naturalmente, la comunidad educativa debe proporcionar resultados útiles - ciertamente, la utilidad motiva la mayor parte del trabajo en Educación - pero es un error pensar que las aplicaciones directas (desarrollo del curriculum, "la prueba" de que las estrategias instruccionales son efectivas, etc.) sean el asunto primario de la investigación en Educación Matemática.

Sobre qué preguntas pueden responderse

Un asunto fundamental que es preciso afrontar al pensar sobre lo que la Educación Matemática puede ofrecer es, ¿qué tipos de cuestiones puede responder la investigación en Educación Matemática?

Planteada de manera simple, la pregunta más típica sobre Educación que se hacen los matemáticos –"¿Qué funciona?" y "¿Qué enfoque es mejor?" – suele en principio no tener respuesta. La razón es que lo que una persona considera efectivo depende de lo que la persona valora. Antes de que se intente decidir si alguna estrategia instruccional tiene éxito, se deben plantear cuestiones tales como: ¿Qué es lo que se quiere lograr exactamente? ¿Qué comprensiones, para qué estudiantes, bajo qué condiciones, con qué restricciones? Consideremos los siguientes ejemplos.

Una cuestión que se plantean con frecuencia los directores y administradores es, "¿Son igual de buenas las clases numerosas que las pequeñas?" Espero que esté claro que esta cuestión no se puede responder de manera abstracta. Lo satisfecho que se puede estar con las clases grandes depende de las consecuencias que se piensa son importantes. ¿Cuánto importa el grado de compromiso de los estudiantes? ¿Son importantes los sentimientos de los estudiantes hacia el curso y hacia el departamento? ¿Interesa el porcentaje de estudiantes que se pueda matricular en cursos posteriores de Matemáticas? Las conclusiones que se pueden extraer respecto de la utilidad de las clases grandes puede variar sustancialmente, dependiendo de la importancia que se conceda a estos resultados.

Preguntas similares surgen incluso si nos centramos sólo en la Matemática que se enseña. Supongamos que se plantea la cuestión, ¿Aprenden los estudiantes tantas matemáticas en las clases grandes como en las clases pequeñas? Se debe preguntar inmediatamente, "¿Qué es lo que cuenta como matemáticas? ¿Qué peso se concederá, por ejemplo, a la resolución de problemas, a la modelización, o a la capacidad para comunicarse matemáticamente?" Los juicios sobre la efectividad de una forma de instrucción sobre otra dependerán de las respuestas que se dé a estas cuestiones. Hablando de manera directa, un investigador tiene que conocer qué buscar y qué tomar como evidencia de lo que se busca antes de ser capaz de determinar si lo ha encontrado.

El que los juicios de una persona reflejan sus valores también se aplica a interrogantes del tipo, ¿Qué estrategia funciona mejor (o es la mejor)? Esto puede parecer obvio, pero con frecuencia no lo es. Consideremos la reforma de la enseñanza del cálculo. Inmediatamente tras el congreso de Tulane *Lean and Lively*, cuyas actas fueron publicadas en Douglas [5], la National Science Foundation (NSF) puso en marcha una importante iniciativa de reforma del cálculo. Hacia mitad de los años noventa los encargados del programa de la NSF estaban convencidos de que la reforma del cálculo era "buena" y que debería ser un modelo para la reforma en otras áreas de contenido. La NSF reunió a los matemáticos que habían estado implicados en la reforma con investigadores en Educación Matemática y planteó la siguiente pregunta: "¿Podemos conseguir evidencia de que la reforma del cálculo funciona (o sea, que la reforma del cálculo es mejor que el cálculo tradicional)?" Lo que tenían en mente, básicamente, era alguna forma de test. Pensaron que sería fácil construir un test, administrarlo, y demostrar que los estudiantes de la reforma lo harían mejor.

Los que eran partidarios de este enfoque no comprendían que lo que proponían sería en esencia como una comparación de manzanas con naranjas. Si se da un test tradicional que se apoya fuertemente en la habilidad para realizar manipulaciones simbólicas, los estudiantes de la "reforma" estarían en desventaja porque no practicaron destrezas de cálculo. Si se da un test que requiera el uso de tecnología o que tenga una fuerte componente de modelización, los estudiantes tradicionales estarían en desventaja porque la tecnología y la modelización no habían sido una parte importante de su currículo. De cualquiera de las dos formas, dar un test y comparar puntuaciones sería injusto. La manera apropiada de proceder consistía en analizar el currículo, identificar tópicos importantes y especificar qué quiere decir tener una comprensión conceptual de esos temas. Con esta clase de información, las instituciones individuales y los departamentos (y la profesión como un todo, si se desea) podría decidir entonces qué aspectos de la comprensión eran más importantes, qué desean evaluar y cómo. Como resultado de amplias discusiones, el esfuerzo del NSF evolucionó desde una posición centrada en documentar los efectos de la reforma del cálculo a otra centrada en desarrollar un marco para observar los efectos de la instrucción sobre el cálculo. El resultado de estos esfuerzos fue el libro de 1997 *Student Assessment in Calculus* [10].

En resumen, muchas de las cuestiones que parecería natural preguntar - cuestiones del tipo, ¿Qué funciona? o ¿Qué método funciona mejor? – no se pueden responder, como se acaba de razonar.

Teniendo en cuenta lo que acabamos de decir, ¿qué tipos de cuestiones se puede plantear la investigación en Educación Matemática? Me propongo argumentar que algunas de las contribuciones fundamentales de la investigación en Educación Matemática son las siguientes:

- perspectivas teóricas para comprender el pensamiento, el aprendizaje y la enseñanza;
- descripciones de aspectos de la cognición (por ejemplo, pensar matemáticamente; lo que los estudiantes comprenden o dejan de comprender sobre los conceptos de función, límite, etc.);
- pruebas de existencia (evidencia de casos en los que los estudiantes pueden aprender a resolver problemas, la inducción, la teoría de grupos; evidencia de la viabilidad de diversos tipos de instrucción);
- descripciones de consecuencias (positivas y negativas) de varias formas de instrucción.

El artículo reciente de Michèle Artigue en *Notices of the American Mathematical Society* [1] describe muchos resultados de tales estudios. Describiré algunos otros y comentaré sobre los métodos de obtenerlos en la siguiente sección sobre "Métodos" más adelante.

Sobre Teorías y Modelos (y Criterios de Calidad)

Cuando los matemáticos usan los términos "teoría" y "modelos" tienen en mente típicamente un serie de cosas muy específicas, tanto sobre la naturaleza de dichas entidades y los tipos de evidencias usadas para hacer afirmaciones sobre ellas. Los términos "teoría" y "modelos" se usan a veces de manera diferente en las ciencias de la vida y en las ciencias sociales, y estos usos pueden ser más similares a los usados en la Educación. En esta sección utilizaré brevemente los ejemplos indicados en la Tabla 1.

Tema	Matemáticas, Física	Biología	Educación, Psicología
Teoría de ...	Ecuaciones, Gravedad	Evolución	Mente
Modelo de ...	Flujo del calor en una placa	Relaciones depredador-presa	Resolución de problemas

Tabla 1.[1] Teorías y modelos en Matemáticas/Física, Biología y Educación/Psicología

En Matemáticas, las teorías se formulan explícitamente, como la teoría de las ecuaciones o la teoría de variables complejas. Los resultados se obtienen analíticamente: probamos que los objetos en cuestión tienen las propiedades que enunciamos sobre ellas. En la Física clásica hay un grado comparable de especificidad; los físicos especifican, por ejemplo, una relación inversa al cuadrado en la atracción gravitatoria. Se entiende que los modelos son aproximaciones, pero se espera que sean aproximaciones muy precisas de forma

[1]Reproducida, con permiso, de [11] página 9.

determinista. Así, por ejemplo, en el modelo del flujo del calor en una placa laminar, especificamos las condiciones iniciales de contorno y las condiciones del flujo de calor, y a continuación resolvemos las ecuaciones relevantes. En suma, no hay ninguna ambigüedad en el proceso. Las descripciones son explícitas, y el estándar de corrección es la prueba matemática. Una teoría y los modelos derivados de ella se pueden usar para hacer predicciones, que a su vez son tomados como verificaciones empíricas de la corrección de la teoría.

Las cosas son bastante más complejas en las Ciencias Biológicas. Consideremos la teoría de la evolución, por ejemplo. Los biólogos están en general de acuerdo sobre su corrección esencial, pero la evidencia acumulada en favor de la evolución es bastante distinta del tipo de evidencia usada en Matemática o en Física. No hay ninguna manera de probar que la evolución es correcta en el sentido matemático; los argumentos que la apoyan consisten en (usando el título de uno de los libros de Pólya) "patrones de razonamiento plausible", junto con la consideración cuidadosa de hipótesis alternativas. En efecto, los biólogos han dicho lo siguiente: "Tenemos montañas de evidencias que son consistentes con la teoría, interpretada de manera amplia; no hay ninguna clara evidencia que contradiga la teoría, y ninguna hipótesis rival cumple los mismos criterios". Aunque la predicción de sucesos futuros no es factible dada la escala de tiempo de los acontecimientos evolutivos, la teoría apoya una forma alternativa de predicción. Los registros de fósiles no examinados previamente se deben ajustar a la teoría, de modo que la teoría puede ser usada para describir propiedades que los fósiles, en estratos geológicos particulares, deberían tener o no. Los registros acumulativos se toman como comprobación de la teoría.

En síntesis, la teoría y la evidencia que la apoya puede diferir sustancialmente en las ciencias de la vida respecto de la Matemática y la Física. Igual ocurre con los modelos, o al menos para el grado de precisión esperado de ellos: nadie espera que las poblaciones de animales representadas mediante las ecuaciones del modelo de depredador-presa se ajusten de la misma manera que el flujo del calor en una lámina plana queda representado con los modelos de flujo del calor.

Finalmente, las teorías y los modelos en las ciencias están siempre sujetas a revisión y refinamiento. A pesar de lo gloriosa y maravillosa que fue, la teoría gravitatoria de Newton fue reemplazada por la teoría de la relatividad de Einstein. Consideremos también la teoría nuclear. La teoría de la valencia, basada sobre modelos de electrones que orbitan alrededor del núcleo, permitió predicciones interesantes, tales como la existencia de elementos todavía no descubiertos. Pero los físicos ya no hablan de electrones que orbitan alrededor del núcleo; una única partícula sólida como los electrones ha sido sustituida en la teoría por nubes probabilísticas de electrones. Las teorías evolucionan.

La investigación en Educación Matemática tiene muchos de los atributos que acabamos de describir de la investigación en Física y en las Ciencias de la Vida. En una "teoría de la mente", por ejemplo, se hacen ciertas hipótesis sobre la naturaleza de la organización mental –por ejemplo, que hay un cierto tipo de estructuras mentales que funcionan de un modo particular. Una de tales hipótesis es que existen varios tipos de memoria, una de las cuales es la memoria de trabajo o "memoria a corto plazo". De acuerdo con la teoría, el pensamiento tiene lugar usando la memoria de trabajo. Lo que hace estas cosas interesantes (y científicas) es que la teoría atribuye a la memoria de trabajo límites bastante fuertes: se afirma (p.e., en [8]) que las personas pueden mantener a lo sumo unas nueve "unidades" de información en la memoria de trabajo al mismo tiempo.

Para ver que esta afirmación puede ser de hecho cierta, uno puede tratar de multiplicar 379 por 658 con los ojos cerrados. La mayoría de las personas encontrarán esta tarea difícil si no imposible. (En una reunión reciente propuse esta tarea a un grupo de setenta y cinco matemáticos. Ninguno de ellos fue capaz de hacerlo en unos pocos minutos). La razón es que el número de cosas que una persona tiene que retener –los números originales y los distintos subtotales que se obtienen al hacer la multiplicación– excede de nueve. Ahora bien, una persona está más capacitada para hacer la tarea mentalmente después de reagrupar algunos de los subtotales: p. e., una persona puede calcular $8 \times 379 = 3032$ y repetir "3032" mentalmente hasta que se convierte en una única agrupación y ocupe una única posición (un "buffer") de la memoria de trabajo. Esto libera suficiente espacio de trabajo para hacer otros cálculos. Usando este tipo de agrupamiento, las personas pueden superar los límites de la memoria de trabajo[2].

Consideremos ahora el estatus de verdad de la afirmación de que la memoria de trabajo de las personas no tiene sólo nueve posiciones. Nunca habrá una prueba absoluta de esta afirmación. En primer lugar, es improbable que los investigadores encuentren la localización física de las posiciones de la memoria de trabajo en el cerebro en el caso de que existan; las posiciones (buffers) son componentes de los modelos, y no son necesariamente objetos físicos. En segundo lugar, la evidencia a favor de esta afirmación es sólida pero no puede ser definitiva. Se han realizado muchos experimentos en los que se dan tareas a las personas que requieren usar más de nueve posiciones en la memoria de trabajo, y las personas han fallado en dichas tareas (o bien, después de algún esfuerzo, las han realizado haciendo lo que se puede considerar como una forma de agrupamiento).

[2] Todo el mundo usa el "agrupamiento" de memoria. Un ejemplo trivial: podemos recordar números de teléfono de nueve cifras memorizando los códigos provinciales como unidades de memoria. La teoría afirma, y esto es lo importante, que el agrupamiento es el mecanismo básico que permite leer este artículo. Cada una de las palabras que uno lee es una unidad, mientras que antes había sido una lista de letras que era necesario oír. Lo mismo ocurre con todo tipo de conceptos matemáticos que manejamos en muestras mentes como unidades. Finalmente, ¿son los calculistas, capaces de realizar computaciones extraordinarias a velocidad de vértigo, un contraejemplo a lo que estamos afirmando? Parece que no. Se deduce de los estudios sobre estos calculistas que memorizan una ingente cantidad de resultados intermedios. Por ejemplo, casi todo el mundo asocia "72" como una unidad al trabajar en un cálculo que incluye 9×8; los calculistas hacen lo mismo con productos de dos y tres cifras. Esto reduce la carga de la memoria de trabajo.

Como ocurre con la evolución, hay muchas evidencias que son consistentes con esta afirmación, no hay ninguna evidencia que la contradiga, y ninguna hipótesis rival satisface los mismos criterios. El estándar relevante es, en esencia, lo que un jurado neutral consideraría como una evidencia fuera de una duda razonable. Lo mismo ocurre para los modelos, por ejemplo, sobre resolución de problemas o sobre (mi interés actual) los modelos de enseñanza (véase [12], [13]). En la actualidad me ocupo de intentar construir una descripción teórica que explique cómo y porqué los profesores hacen lo que hacen, sobre la marcha, en clase. Este trabajo, elaborado al mismo nivel de detalle que la teoría de la memoria, es llamado una "teoría de la enseñanza-en-contexto". Lo que se afirma es que con la teoría y con tiempo suficiente para modelizar a un profesor en concreto, se puede construir una descripción de la enseñanza de esa persona que caracteriza su comportamiento en la clase con notable precisión. Cuando miramos a este trabajo, no se puede esperar encontrar el tipo de precisión que se encuentra en la modelización del flujo de calor en una lámina plana. Sin embargo, es razonable esperar (véase, p.e., [12]) que tal conducta pueda modelizarse con el mismo grado de fidelidad respecto del comportamiento en el "mundo real" que en los modelos de depredador-presa.

Continuaremos con la cuestión de los estándares de evaluación de las teorías, modelos y resultados en la sección que sigue después de la próxima.

Métodos

En este artículo no puedo proporcionar siquiera un primer catálogo de métodos de investigación en Educación Matemática universitaria. ¡Como una indicación de la magnitud de esa tarea consideren el hecho de que el *Handbook of Qualitative Research in Education* [6] tiene casi 900 páginas! Los capítulos de ese volumen incluyen extensas discusiones de etnografía (¿cómo se entiende la "cultura de la clase", por ejemplo?), análisis del discurso (¿qué patrones se pueden ver en el estudio cuidadoso de las conversaciones?), el papel de la cultura en la conformación de la cognición, y cuestiones de subjetividad y validez. Y esto se refiere sólo al trabajo de tipo cualitativo-naturalmente, existe también una larga tradición de investigación cuantitativa en las ciencias sociales. Mi propósito es más bien proporcionar una orientación de que tipo de trabajos se están haciendo y sugerir que descubrimientos (y limitaciones de los mismos) a que pueden dar lugar.

Los que se acercan por primera vez a la investigación educativa tienden a pensar en términos de estudios experimentales tradicionales, que implican grupos experimentales y de control y el uso de la estadística para determinar si los resultados son o no significativos. Sin embargo, el uso de la estadística en la educación es una cuestión mucho más complicada de lo que se podría pensar.

Desde hace algunos años, a partir de la mitad de este siglo, la investigación en las ciencias sociales (por lo menos en los Estados Unidos) estuvo dominada por el ejemplo de la agricultura. La noción básica era que si dos campos de una cosecha particular fueron tratados de manera idéntica excepto en una variable, entonces las diferencias en las cosechas producidas se podrían atribuir a la diferencia en esa variable. Seguramente, la gente pensó, que se podría hacer lo mismo en Educación. Si se quiere probar que un nuevo método de enseñanza del contenido X es superior, se podría hacer un experimento en el que dos grupos de alumnos estudian X –un grupo recibe una enseñanza tradicional, mientras que al otro se le enseña con el nuevo método. Si los estudiantes que recibieron la nueva enseñanza puntúan mejor, se tendría evidencia de la superioridad del nuevo método instruccional.

Apartemos de momento las cuestiones planteadas en la sección anterior sobre los fines de la instrucción y el hecho de que la instrucción antigua y la nueva podrían no centrarse en las mismas cosas. Imaginemos que se puede construir un test equitativo para comparar estas formas de instruir. Y supongamos que los estudiantes fueron asignados aleatoriamente a los grupos experimentales y de control, y que se aplicaron procedimientos de experimentación tradicionales. A pesar de todo esto, todavía podría haber problemas potencialmente graves. Si a los dos grupos de estudiantes les enseñan diferentes profesores, cualquier diferencia en los resultados podría atribuirse a las diferencias en la enseñanza. Pero incluso con el mismo profesor, podría haber innumerables diferencias. Puede haber una diferencia en la energía o compromiso: enseñar el “mismo viejo material” no es lo mismo que intentar nuevas ideas. O los estudiantes en uno de los grupos pueden saber que están probando algo nuevo y experimental. Sólo esto puede provocar diferencias significativas. (Hay una extensa bibliografía que muestra que si sienten que se están haciendo cambios por su propio interés, trabajarán más duro y mejor –sin importar de hecho cómo sean los cambios. Los efectos de estos cambios se desvanecen con el tiempo.) O bien los estudiantes se pueden retraer si sienten que se experimenta con ellos.

He aquí un ejemplo sencillo en este sentido. Hace algunos años desarrollé unos materiales instruccionales para el cálculo. Colegas de otra universidad lo adoptaron también para sus alumnos. En todas las secciones excepto en dos los estudiantes que los usaron puntuaron mejor que los estudiantes que no los usaron. Sin embargo, en dos secciones no había esencialmente ninguna diferencia en el rendimiento. Resultó que la mayor parte de los profesores habían dado una introducción favorable a los materiales, sugiriendo a los estudiantes que les serían útiles. El profesor que impartió las dos secciones en las que no se mostró ninguna diferencia los había entregado diciendo, “Me han pedido que les entregue estos materiales, no sé si son buenos”.

En resumen, los métodos experimentales clásicos pueden ser problemáticos en la investigación experimental. Para señalar al menos dos dificultades, los experimentos de doble ciego en el sentido médico (en el que ni los médicos ni los pacientes saben quién está recibiendo el tratamiento real y quién está recibiendo un tratamiento placebo) son raramente ciegos, y muchas variables experimentales son difícilmente controlables en un sentido riguroso. (Ése era el caso del ejemplo del párrafo anterior.) Como resultado de esto, tanto los resultados positivos como los negativos son difíciles de interpretar. Esto no quiere decir que tales estudios no sean útiles o que el trabajo estadístico en gran escala no sea valioso –claramente lo es– sino que debe ser hecho con un gran cuidado y que los resultados y las afirmaciones deben ser interpretadas con igual cuidado. El trabajo estadístico de valor consistente suele ser aquel que

a) produce descubrimientos generales sobre una población. Por ejemplo, Artigue [1] observa que "más del 40% de los estudiantes que entran en las universidades francesas consideran que si la diferencia entre dos números A y B es menor que $1/N$ para todo número positivo N, dichos números no son necesariamente iguales, sino infinitamente próximos".

b) proporciona una clara comparación de dos o más poblaciones. Por ejemplo, los resultados del Tercer Estudio Internacional sobre Matemáticas y Ciencias documentan el rendimiento básico de los estudiantes en diversas naciones sobre un abanico de contenidos matemáticos.

c) Proporciona apoyo, a lo largo del tiempo, de descubrimientos que fueron hechos primeramente en estudios observacionales de menor escala.

Lo que aportan los métodos de investigación en Educación Matemática universitaria –o sobre Educación en cualquier otra materia– es sugerencias de resultados; la evidencia combinada de muchos estudios a lo largo del tiempo es lo que concede apoyo a los descubrimientos.

Me extenderé sobre este punto con un ejemplo extraído de mi propio trabajo. El tema se refiere a la "conducta metacognitiva", o metacognición: específicamente, el uso efectivo de los propios recursos (incluyendo el tiempo) durante la resolución de problemas.

Este es un ejemplo motivador. Hace muchos años, cuando un tema de un curso estándar de cálculo de primer año era las técnicas de integración, el siguiente ejercicio era el primer problema de un test dado a una clase numerosa:

$$\int \frac{x}{x^2 - 9} dx\,.$$

Lo que se esperaba era que los estudiantes hicieran la sustitución obvia $u = (x^2 - 9)$ y resolvieran el problema de manera breve. Aproximadamente la mitad de la clase así lo hizo. Sin embargo, alrededor de la cuarta parte de la clase, al advertir que el denominador era factorizable, intentaron resolver el problema mediante las fracciones simples. Además, como un 10 por ciento de los estudiantes, tras observar que el denominador era de la forma $(x^2 - a^2)$, intentaron resolver el problema usando la sustitución $x = 3 \cdot \mathrm{sen}(\theta)$. Todos estos resultados dan la respuesta correcta, naturalmente, pero el segundo y el tercer método requieren mucho más tiempo. Los estudiantes que usaron esas técnicas puntuaron bajo en la prueba, principalmente porque les faltó tiempo.

Ejemplos como estos me llevaron a desarrollar material instruccional centrado en las elecciones estratégicas que se hacen en cálculo integral. Los materiales lograron una diferencia en el rendimiento de los estudiantes. Esto proporciona alguna evidencia de que las elecciones estratégicas durante la resolución de problemas son importantes.

La cuestión de las elecciones estratégicas apareció de nuevo cuando, como parte de mi investigación sobre resolución de problemas, examiné grabaciones en vídeo de estudiantes intentando resolver problemas. Con frecuencia, parecía como que los estudiantes leían el enunciado del problema, elegían rápidamente un método de solución, y después tenazmente seguían ese enfoque incluso cuando no parecía producir resultados. Para justificar estas observaciones rigurosamente, desarrollé un "esquema de codificación" para analizar las grabaciones de vídeo de la resolución de problemas. Este marco analítico proporcionó un mecanismo para identificar aquellos momentos durante la resolución de un problema en los que la decisión adoptada podía determinar el éxito o el fracaso del intento. El esquema de análisis se definió de forma que otros investigadores pudieran usarlo, no sólo para que pudieran estudiar mis vídeos, sino para que también pudieran analizar los suyos. Con este procedimiento, los investigadores podían ver cómo las decisiones adoptadas por los estudiantes ayudaban o entorpecían sus intentos de resolver los problemas.

Estos esquemas analíticos sirven para diversos propósitos. En primer lugar, el disponer de uno de estos esquemas permite que la caracterización de grabaciones se pueda hacer de una manera relativamente objetiva: si dos analistas entrenados trabajando sobre la misma cinta independientemente producen la misma codificación, entonces hay razón para creer en la consistencia de la interpretación. En segundo lugar, teniendo una herramienta analítica de este tipo podemos seguir los efectos de la instrucción en resolución de problemas: la comparación de grabaciones "antes y después" de las sesiones de resolución de problemas puede revelar si los estudiantes han logrado ser más efectivos o eficientes en la resolución de problemas. En tercer lugar, este tipo de herramientas permite la acumulación de datos entre distintos estudios. Un breve resumen de los resultados en este caso podría ser el siguiente: la competencia metacognitiva es un factor muy productivo en la resolución de problemas[3]. Para detalles más extensos, véase [9].

Como se ha indicado anteriormente, los resultados de la investigación en Educación no se "demuestran" en el sentido en el que se demuestra en Matemáticas. Además, con frecuencia es difícil emplear métodos experimentales o estadísticos directos del tipo de los usados en las Ciencias Físicas debido a lo complicado que resulta decidir qué se quiere decir con que unas condiciones educativas se han "replicado". En Educación se dispone de una amplia variedad de métodos de investigación. Un vistazo a uno de los primeros volúmenes sobre Educación Matemática universitaria, a saber, [14], sugiere este rango de métodos. El número y tipo de métodos se ha incrementado, como evidencian los tres volúmenes de *Research Collegiate Matematics Education*. Se encuentran, por ejemplo, informes de entrevistas detalladas con estudiantes, comparaciones de cursos de cálculo de la reforma con cursos tradicionales, un

[3]En cuanto al comportamiento metacognitivo, una gran cantidad de estudios han señalado que la capacidad de tomar decisiones efectivas durante la resolución de problemas no es "natural". Es una habilidad que puede aprenderse, aunque se requiere instrucción intensiva. Cuando los estudiantes adquieren tales destrezas, su capacidad para resolver problemas mejora notablemente.

examen de “talleres” sobre cálculo, y un amplio estudio sobre el desarrollo de la comprensión por el estudiante de un dispositivo físico y gráficos relacionados con él. Los estudios que emplean técnicas de observación antropológica y otros métodos cualitativos son cada vez más comunes.

¿Cuál es la validez de tales estudios, y cuánto podemos depender de sus resultados? Esta cuestión se estudia en la siguiente sección.

CRITERIOS PARA JUZGAR TEORÍAS, MODELOS Y RESULTADOS

Hay una gran variedad de resultados y métodos en Educación Matemática. Una pregunta fundamental es la siguiente: ¿Cuánta confianza deberíamos depositar en un resultado concreto? ¿Qué constituye una razón sólida, qué constituye “una prueba libre de duda razonable”?

La siguiente lista propone un conjunto de criterios que pueden usarse para evaluar modelos y teorías y (de manera más general) cualquier trabajo empírico o teórico en Educación Matemática:

- Poder descriptivo
- Poder explicativo
- Alcance
- Poder predictivo
- Rigor y especificidad
- Posibilidad de refutación
- Replicabilidad
- Disponibilidad de múltiples fuentes de evidencia (“triangulación”)

Describiré brevemente cada uno de estos criterios.

Poder Descriptivo. Mediante poder descriptivo me refiero a la capacidad de una teoría para captar “lo que cuenta”, de manera que parezca fiel al fenómeno que se describe. Como Gaea Leinhart [7] ha señalado, la frase “consideremos una vaca esférica” puede ser apropiada cuando los físicos están considerando la vaca en términos de su masa gravitacional –pero no si se está explorando alguna de las propiedades fisiológicas de la vaca. Las teorías de la mente, de la resolución de problemas, o de la enseñanza deberían incluir aspectos relevantes e importantes del pensamiento, la resolución de problemas, y la enseñanza, respectivamente. En un sentido muy amplio, preguntas pertinentes que se deben hacer son: ¿Falta algo? ¿Los elementos de la teoría se corresponden con cosas que parecen razonables? Por ejemplo, una sesión de resolución de problemas, una entrevista, o una lección ha sido grabada en vídeo. ¿Una persona que leyera el análisis y viera el vídeo se puede sorprender razonablemente por cosas que falten en el análisis?

Poder Explicativo. Por poder explicativo quiero decir proporcionar explicaciones de cómo y porqué funcionan las cosas. Una cosa es decir que las personas serán o no capaces de hacer ciertos tipos de tareas o incluso describir detalladamente lo que hacen; otra cosa bastante diferente es explicar por qué. Una cosa es, por ejemplo, decir que las personas tendrán dificultades al multiplicar números de tres cifras de memoria. Pero eso no aporta información sobre cómo y porqué ocurren esas dificultades. La descripción teórica completa de la memoria de trabajo, que se ha mencionado anteriormente, aporta una descripción de las posiciones de memoria, una descripción detallada del mecanismo de agrupación, y la descripción cuidadosa de cómo los componentes de la memoria interactúan unos con otros. La explicación se hace a un nivel de mecanismo: dice en términos razonablemente precisos cuáles son los objetos de la teoría, cómo están relacionados, y por qué algunas cosas serán posibles y otras no.

Alcance. Con alcance me refiero al rango de fenómenos cubiertos por la teoría. Una teoría de ecuaciones no es muy impresionante si trata sólo sobre ecuaciones lineales. De modo semejante, una teoría de la enseñanza no es muy relevante si se aplica sólo a lecciones magistrales.

Poder Predictivo. El papel de la predicción es obvio: una prueba para cualquier teoría es si puede especificar algunos resultados antes de que tengan lugar. De nuevo, es bueno tener en la mente como modelo las características de teorías tales como la teoría de la evolución. Las predicciones en Educación y Psicología no son con frecuencia como las que se hacen en Física.

A veces es posible hacer predicciones precisas. Por ejemplo, Brown y Burton [4] estudiaron qué comprensiones incorrectas adquieren los estudiantes cuando aprenden el algoritmo estándar de la resta en base 10. Hicieron hipótesis sobre construcciones mentales muy específicas de los estudiantes –la idea es que los alumnos simplemente no fallan en dominar el algoritmo, sino más bien que los alumnos a menudo desarrollan una gran variedad de variantes incorrectas del algoritmo y las aplican sistemáticamente. Brown y Burton desarrollaron un test diagnóstico simple con la propiedad de que el patrón de respuesta incorrecta de un alumno indicaba el algoritmo falso que podía estar usando. En la mitad del tiempo eran capaces de predecir la respuesta incorrecta que el estudiante daría a un problema nuevo, ¡antes de que el estudiante hiciera el problema!

Este tipo de predicciones consistentes y detalladas sobre la base de algo tan simple como una prueba de diagnóstico son, por supuesto, extremadamente raras. Por ejemplo, ninguna teoría sobre la enseñanza puede predecir de modo preciso lo que un profesor hará en diversas circunstancias; la conducta humana no se puede predecir de esa manera. Sin embargo, una teoría de la enseñanza puede funcionar de manera análoga a como lo hace la teoría de la evolución. Puede sugerir restricciones e incluso sugerir sucesos probables.

[Hacer predicciones es una herramienta muy potente en el refinamiento de la teoría. Cuando algo se afirma que es imposible y sucede, o cuando una teoría hace afirmaciones repetidas de que algo es muy probable y no ocurre, ¡entonces la teoría tiene problemas! Así pues, comprometerse en tales predicciones es una herramienta metodológica importante, incluso cuando se comprende que las predicciones precisas son imposibles.]

Rigor y Especificidad. La construcción de una teoría o un modelo implica la especificación de un conjunto de objetos y de relaciones entre ellos. Este conjunto de objetos abstractos y de relaciones supuestamente se corresponde con algún conjunto de objetos y relaciones en el "mundo real". Las cuestiones relevantes son:

¿Están bien definidos los términos? ¿Se identificará un objeto si se ve? ¿En la vida real? ¿En el modelo? ¿Están bien definidas las relaciones entre los objetos? ¿Cómo de bien se corresponden los objetos y relaciones en el modelo con las cosas que se supone representan? Como se ha observado anteriormente, no podemos esperar necesariamente el mismo tipo de correspondencia entre las partes del modelo y los objetos del mundo real como en el caso de los modelos físicos simples. Constructos mentales y sociales tales como posiciones de memoria o el "contrato didáctico" (la idea de que los profesores y los estudiantes entran a la clase con acuerdos implícitos sobre las normas que regulan sus interacciones y que esas comprensiones conforman los modos en que actúan) no son examinables o medibles de la misma manera que el flujo de calor en una lámina plana. Pero podemos exigir detalles, tanto sobre lo que los objetos son y cómo se ajustan. ¿Se definen las relaciones y los cambios entre ellos cuidadosamente, o "aparecen de manera mágica" en algún lugar y en algún momento? A continuación describo una analogía aproximada. Durante la mayor parte del siglo XVIII la teoría de la combustión del flogisto –que postulaba que en todos los materiales inflamables hay una sustancia llamada "flogisto" que era incolora, inodora, sin peso ni sabor que se libera durante la combustión –fue ampliamente aceptada. (El trabajo de Lavoisier sobre la combustión refutó finalmente la teoría). Como por arte de magia, la teoría del flogisto explicaba una variedad razonable de fenómenos. Se podría haber continuado usándola, de modo similar a como los teóricos podrían haber continuado construyendo epiciclos sobre más epiciclos en una teoría de las órbitas circulares[4]. La teoría puede seguir produciendo algunos resultados útiles, suficientemente buenos "con fines prácticos". Esto puede ser suficiente para la práctica, pero no para la teoría. Como ocurre en las Ciencias Físicas, los investigadores sobre Educación tienen la obligación intelectual de lograr la mayor

[4]Este ejemplo indica otro criterio importante, simplicidad. Cuando una teoría requiere múltiples "ajustes" tales como epiciclos sobre más epiciclos, esto es un síntoma de que algo no va bien.

claridad y especificidad e identificar los casos límite o contraejemplos para ver donde fallan las ideas teóricas.

Muestro a continuación dos rápidos ejemplos. Primero, en el modelo del proceso de enseñanza desarrollado por mi equipo de investigación representamos aspectos tales como los conocimientos, metas, creencias y toma de decisiones del profesor. Los escépticos (entre los que nos incluimos) preguntarían, ¿cómo de clara es la representación? Una vez que se definen los términos en el modelo (p.e., cuando especificamos un conocimiento del profesor, fines y creencias), ¿se levanta la mano cuando decimos lo que el profesor puede hacer en circunstancias específicas?, o ¿es el modelo lo suficientemente bien definido como para que otras personas puedan aplicarlo y hacer las mismas predicciones? Segundo, "la teoría APOS" presentada en [2] usa términos tales como Acción, Proceso, Objeto, y Esquema. ¿Reconoceríamos estas nociones si nos encontrásemos con ellas? ¿Está bien especificado cómo interactúan o se transforman? En ambos casos las cuestiones básicas son, ¿Cuáles son las posibilidades de que esta teoría sea semejante a la teoría del flogisto? ¿Las personas que emplean la teoría están constantemente probándola con el fin de contrastarla? Cuestiones similares se deberían preguntar respecto de todos los términos que se usan en la investigación educativa, p.e., "el contrato didáctico", "metacognición", "imagen de un concepto", y "obstáculo epistemológico".

Posibilidad de refutación. La necesidad de falsación –de hacer afirmaciones no tautológicas o predicciones cuya precisión pueda ser probada empíricamente –debería estar ya clara a estas alturas, pues es concomitante con la discusión hecha en los dos apartados previos. Un campo hace progresos (y se protege de las tautologías) poniendo sus ideas a prueba.

Replicabilidad. La cuestión de la replicabilidad está también íntimamente ligada a la del rigor y la especificidad. Hay dos preguntas relacionados:

1. ¿Ocurrirán las mismas cosas bajo las mismas circunstancias?
2. ¿Otras personas, una vez entrenadas adecuadamente, observarán los mismos hechos en los datos?

En ambos casos la respuesta a estas cuestiones depende de tener bien definidos los procedimientos y los constructos.

El enunciado de (1) es deliberadamente vago, porque quiere cubrir un amplio abanico de casos. En el caso de la memoria a corto plazo, la afirmación es que las personas tendrán dificultades si las tareas de memoria requieren el uso de más de nueve posiciones de memoria a corto plazo. En el caso de los análisis sociológicos del aula, la afirmación es que una vez que se comprende el contrato didáctico, las acciones de los estudiantes y del profesor se verá que se ajustan a dichas reglas (usualmente tácitas). En el caso de las creencias, la afirmación es que los estudiantes que mantienen ciertas creencias actuarán de cierto modo cuando hacen matemáticas. En el caso de los obstáculos epistemológicos o la teoría APOS, se hacen afirmaciones similares de que los estudiantes que tienen (o no tienen) ciertas concepciones mentales particulares serán capaces de hacer (o de no hacer) ciertas cosas.

En todos estos casos la utilidad de los descubrimientos, la precisión de las afirmaciones, y la habilidad para refutar o replicar depende de la especificidad con la que se definen los términos. Consideremos el siguiente ejemplo extraído de la bibliografía clásica sobre Educación. La teoría de Ausubel de los "organizadores avanzados" (véase [3]) postula que si se da a los estudiantes una introducción a los materiales que tienen que leer que les oriente respecto de lo que sigue, su comprensión de la lectura mejorará significativamente. Después de una década o dos de muchos estudios, la bibliografía sobre el tema no alcanzó ninguna conclusión: cerca de la mitad de los estudiantes mostraron que los organizadores avanzados producían una diferencia, mientras que en la otra mitad no se tenía efecto. Un análisis más fino reveló la razón: los propios términos estaban mal definidos. Varios investigadores han construido sus propios organizadores avanzados basados en lo que ellos pensaban que deberían de ser –y hubo una gran variación. ¡No es de extrañar entonces que las conclusiones no fueran concluyentes! (Una técnica tradicional para tratar cuestiones relativas a las buenas definiciones de los constructos, y que afronta la cuestión (2) anterior, consiste en tener investigadores independientes trabajando sobre el mismo cuerpo de datos y comparar sus resultados. Hay normas estándar en el campo para la "fiabilidad inter-evaluadores"; estas normas cuantifican el grado en el que analistas independientes ven las mismas cosas en los datos).

Múltiples Fuentes de Evidencia ("Triangulación"). Aquí encontramos una de las principales diferencias entre las Matemáticas y las Ciencias Sociales. En Matemáticas una argumentación convincente (una prueba) es suficiente: esto establece la validez. En Educación y en las Ciencias Sociales una tarea fundamental es la búsqueda de evidencias convincentes. La realidad es que la evidencia puede ser engañosa: lo que pensamos que es general puede de hecho ser algo artificial o función de las circunstancias más que un fenómeno general.

Veamos el siguiente ejemplo. Hace algunos años grabé en vídeo a estudiantes universitarios que discurrían sobre la pregunta: ¿Cuántas células hay, en promedio, en el cuerpo de un humano adulto? Sus conductas fueron llamativas. Una cierta cantidad de estudiantes hicieron estimaciones salvajes del orden de magnitud de las dimensiones de una célula -desde "supongamos que una célula tiene un angstrom de largo" a "supongamos que una célula es un cubo de 1/100 de pulgada de lado". A continuación, una vez que decidieron en unos pocos segundos el tamaño de la célula, dedicaron un largo tiempo a determinar el tamaño del cuerpo, descomponiendo a menudo el cuerpo en una colección de cilindros, conos, y esferas y calculando el volumen de cada uno con cuidado. Esto era *muy* extraño.

Algún tiempo después comencé a grabar a los estudiantes resolviendo los problemas por parejas en lugar de individualmente. Nunca observé el tipo de conducta que acabo de describir. Resultó que cuando trabajaban solos, los estudiantes se sentían bajo una presión tremenda. Sabían que un profesor de

Matemáticas estaría mirando su trabajo. Bajo estas circunstancias sentían que necesitaban hacer algo matemático, ¡y los cálculos de volúmenes al menos hacía parecer como si estuvieran haciendo matemáticas! Cuando los estudiantes trabajan en parejas, comenzaban diciendo algo como, “Este es un problema ciertamente raro”. Esto era suficiente para disminuir algo la tensión, el resultado era que no había necesidad de comenzar a hacer cálculos de volúmenes para mitigar la tensión. En resumen, cierta conducta muy consistente era de hecho función de las circunstancias en lugar de ser inherente al problema o a los estudiantes.

Una manera de comprobar la existencia de conductas artificiales consiste en variar las circunstancias: preguntar, ¿ves la misma cosa en diferentes momentos y en lugares distintos? Otra manera consiste en buscar tantas fuentes de información como sea posible sobre el fenómeno en cuestión y ver si representan un mensaje consistente. En el trabajo sobre modelización de la enseñanza de mi grupo de investigación, por ejemplo, extraemos inferencias sobre la conducta del profesor a partir de grabaciones en vídeo del profesor en acción –pero también entrevistamos al profesor, revisamos las programaciones de sus lecciones y notas de clase, y discutimos nuestras posibles averiguaciones con el propio profesor. De este modo buscamos convergencia de datos. Cuantas más fuentes independientes de confirmación haya, más robusto será probablemente un descubrimiento.

Conclusión

El tema principal de este artículo ha sido que la investigación sobre Educación Matemática (en la Universidad) es una empresa muy diferente de la investigación en Matemáticas, y que la comprensión de las diferencias es esencial para poder apreciar el trabajo en este campo (o mejor aún, contribuir a dicho trabajo). Los descubrimientos son raramente definitivos; con frecuencia son simplemente sugerencias. La evidencia no es del tipo de las demostraciones, sino que es acumulativa, progresando hacia conclusiones que se pueden considerar como fuera de toda duda razonable. Una aproximación científica es posible, pero se debe tener cuidado para no ser cientifista –lo que cuenta no son los adornos de la ciencia, tales como el método experimental, sino el uso del razonamiento cuidadoso y los estándares de evidencia, empleando una amplia variedad de métodos apropiados para la tarea correspondiente.

Debemos recordar lo reciente que es la Educación Matemática como campo de investigación. Los matemáticos están acostumbrados a medir el linaje matemático en siglos, cuando no en milenios; en contraste, el linaje de la investigación en Educación Matemática (especialmente la Educación Matemática universitaria) se mide en décadas. La revista *Educational Studies in Mathematics* comenzó a editarse en los años 60. El primer número del Volumen 1 del *Journal for Research in Mathematics Education* fue publicado en Enero de 1970. Las series de volúmenes de *Research in Collegiate Mathematics Education* –el primer conjunto de volúmenes dedicado exclusivamente a Educación Matemática en el nivel universitario –comenzó a publicarse en 1994. No es una casualidad que la gran mayoría de los artículos citados por Artigue [1] en su revisión de 1999 sobre contribuciones de la investigación fueron escritos en los años 90; ¡antes de esa fecha había poco dedicado al nivel universitario! Ha habido un progreso extraordinario en los años recientes, pero el campo es todavía muy reciente, y queda mucho camino por recorrer.

Debido a la naturaleza del campo, es apropiado manifestar la propia posición hacia el trabajo y su utilidad. Los matemáticos que se aproximan a este trabajo deberían estar abiertos a una amplia variedad de ideas, comprendiendo que los métodos y perspectivas a los que están acostumbrados no se aplican a la investigación educativa de manera directa. No deberían buscar respuestas definitivas sino ideas que se pueden usar. Al mismo tiempo, todos los usuarios de la investigación en Educación Matemática universitaria deberían ser saludablemente escépticos. En particular, debido a que no hay respuestas definitivas, ciertamente se debería desconfiar de alguien que las ofrece. En general, el principal objetivo para las próximas décadas consiste en continuar construyendo un cuerpo de teoría y métodos que permita a la investigación en Educación Matemática llegar a ser un campo cada vez más sólido, tanto básico como aplicado.

Referencias

[1] M. Artigue, *The teaching and learning of mathematics at the university level: Crucial questions for contemporary research in education,* Notices Amer. Math. Soc. 46 (1999), 1377-1385.

[2] M. Asiala, A. Brown, D. De Vries, E. Dubinsky, D. Mathews, y K. Thomas, *A framework for research and curriculum development in undergraduate mathematics education, Research in Collegiate Mathematics Education* (J. Kaput, A. Schoenfeld, y E. Dubinsky, editores.), volumen II, Conference Board of the Mathematical Sciences, Washington, DC, pp. 1-32.

[3] D. P. Ausubel, *Educational Psychology: A Cognitive View,* Holt-Reinhart-Winston, New York, 1968.

[4] J. S. Brown y R. R. Burton, *Diagnostic models for procedural bugs in basic mathematical skills*, Cognitive Science 2 (1978), 155-192.

[5] R. G. Douglas (editor), *Toward a Lean and Lively Calculus*, MAA Notes Number 6, Mathematical Association of America, Washington, DC, 1986.

[6] M. Lecompte, W. Millroy, y J. Preissle, *Handbook of Qualitative Research in Education,* Academic Press, New York, 1992.

[7] G. Leinhardt, *On the messiness of overlapping goals in real settings,* Issues in Education 4 (1998), 125-132.

[8] G. Miller, *The magic number seven, plus or minus two: Some limits on our capacity for processing information,* Psychological Review 63 (1956), 81-97.

[9] A. H. Schoenfeld, *Mathematical Problem Solving,* Academic Press, Orlando, FL, 1985.

[10] A. H. Schoenfeld, (editor), *Student Assessment in Calculus,* MAA Notes Number 43, Mathematical Association of America, Washington, DC, 1997.

[11] A. H. SCHOENFELD, *On theory and models: The case of teaching-in-context,* Proceeding of the XX Annual Meeting of the International Group for Psychology and Mathematics Education (Sarah B. Berenson, ed.), Psychology and Mathematics Education, Raleigh, NC, 1998.
[12] A. H. SCHOENFELD, *Toward a theory of teaching-in-context,* Issues in Education 4 (1998), 1-94.
[13] A. H. SCHOENFELD, *Models of the teaching process,* Journal of Mathematical Behavior (in press).
[14] D. TALL (editor), Advanced Mathematical Thinking, Kluwer, Dordrech, 1991.

Alan H. Schoenfeld
Elizabeth and Edward Conner Professor of Education
Universidad de California, Berkeley, Estados Unidos
correo electrónico: alan@socrates.berkeley.edu.

Traducción de Juan D. Godino, Universidad de Granada

Algunas ideas y lecturas complementarias al artículo de Schoenfeld

por

Juan D. Godino

El artículo de Alan H. Schoenfeld aporta una visión de la investigación en Educación Matemática que presenta, de manera nítida, su naturaleza diferente, respecto de la investigación en Matemáticas, y sus semejanzas respecto a la investigación en las ciencias experimentales y sociales. Es importante la distinción que se hace en el trabajo entre los dos propósitos principales de la investigación en educación matemática, uno puro (propio de la ciencia básica) y otro aplicado (característico de la ingeniería), y las relaciones entre ambos ámbitos de indagación. También queda clara la variedad y complejidad de los métodos de investigación aplicables, así como los diversos criterios para evaluar la calidad y relevancia de las investigaciones.

El autor utiliza las citas iniciales de Cohen y Pollak como apoyo para resaltar "algunas de las principales diferencias entre la matemática y la educación matemática –diferencias que se deben entender si se desea comprender la naturaleza de los métodos y resultados en la educación matemática". También afirma en las conclusiones que el "tema principal de este artículo ha sido que la investigación sobre Educación Matemática es una empresa muy diferente de la investigación en Matemáticas".

En mis observaciones a este trabajo no voy a negar la existencia y pertinencia de estas diferencias, pero considero importante también tener en cuenta algunas semejanzas entre el trabajo del matemático que investiga y construye nuevos conocimientos y el investigador en Educación Matemática ("didacta"), al menos dentro de la perspectiva de la didáctica de la matemática que se describe como "aproximación epistemológica" (Gascón, 1998; Brousseau, 1997). En este enfoque de investigación se entra al estudio de los problemas de la enseñanza y aprendizaje de las matemáticas a través del "polo del saber matemático", esto es, mediante la búsqueda y caracterización de las situaciones matemáticas que dan sentido a los contenidos que se proponen enseñar y el cuestionamiento epistemológico del propio conocimiento matemático.

Esta búsqueda de problemas, en el desarrollo histórico de la matemática y en la fenomenología matemática, tiene unas características propias de la actividad matemática, aunque ciertamente que la indagación de la viabilidad de una propuesta de ingeniería didáctica tiene que ser abordada con las herramientas y diseños propios de las ciencias experimentales y sociales.

Una parte importante de la investigación sobre el componente científico de la educación matemática se debe concebir, como propone Brousseau, como una "epistemología experimental" de las matemáticas, debiendo afrontar, en consecuencia, delicadas cuestiones de tipo ontológico, epistemológico, semiótico, además, de las dimensiones psicológicas y socioculturales. De esta manera, el contexto de la Educación Matemática, proporciona un nuevo ámbito de reflexión e indagación a la epistemología de la matemática –entendida en sentido ampliado–, que se añade a los contextos de invención-descubrimiento, justificación y aplicación. El estudio de las adaptaciones de los conocimientos matemáticos para ser enseñados en los distintos contextos institucionales, así como la búsqueda de criterios para su selección y articulación es una parte esencial de la problemática didáctica no contemplada en la lista de cuestiones citadas en el artículo de Schoenfeld. La indagación de estos problemas de transposición didáctica y ecología de los saberes matemáticos tiene unas características semejantes al trabajo del matemático y deben ser abordados dentro de la didáctica de las matemáticas.

Por otra parte, el trabajo de reorganización y reformulación del saber matemático es una parte importante de la actividad del productor de matemáticas; esa misma actividad tiene que ser realizada por el investigador en didáctica de las matemáticas a fin de que los estudiantes puedan proceder a la reconstrucción de los conocimientos matemáticos. Las "semejanzas" entre el trabajo del matemático y el del didacta no se refieren a la naturaleza de los conocimientos producidos en cada caso, los procedimientos de validación y organización de tales conocimientos, sino a la actividad de búsqueda y solución de problemas matemáticos.

Finalmente, considero que el artículo ofrece una visión lógico-deductiva, de las matemáticas, que es algo parcial: "En matemáticas, [. . .] los resultados se obtienen analíticamente: probamos que los objetos en cuestión tienen las propiedades que enunciamos sobre ellas". Aunque esto es cierto si nos atenemos a la organización final de los contenidos matemáticos, el proceso de creación matemática es mucho más complejo, ya que requiere poner en juego todos los recursos intelectuales de la persona. Igualmente ocurre con los procesos de reinvención guiada (Freudenthal, 1991) que se requiere promover en el seno de la clase de matemáticas y cuyo estudio debe abordar la didáctica de las matemáticas.

En síntesis, cuando adoptamos un modelo epistemológico apropiado sobre la actividad matemática y sus producciones culturales, la investigación sobre una parte importante de los problemas de enseñanza y aprendizaje de las matemáticas adquiere connotaciones propias de la investigación matemática, no en cuanto a la organización deductiva de los resultados matemáticos, sino en lo referente a los procesos de reinvención y descubrimiento que se ponen en juego en ambas disciplinas. Si atendemos a estos procesos, la didáctica de la matemática se relaciona estrechamente con la actividad matemática, pudiendo aportar descripciones y explicaciones del propio desarrollo de la matemática, concebida como una construcción humana.

Referencias y lecturas complementarias recomendadas

Brousseau, G. (1989). ¿Qué puede aportar la didáctica de las matemáticas a un profesor? *Suma*, 4, 5-12 y *Suma*, 5, 5-12.

Brousseau, G. (1997). *Theory of didactic situations in mathematics.* Dordrecht: Kluwer A. P.

Freudenthal, H. (1991). *Revisiting mathematics education.* Dordrecht: Kluwer A. P.

Gascón, J. (1998). Evolución de la didáctica de las matemáticas como disciplina científica. *Recherches en Didactique des Mathématiques,* 18(1): 7-34.

Godino, J. D. (1991). Hacia una teoría de la didáctica de las matemáticas. En A. Gutierrez (Ed.), *Área de conocimiento: Didáctica de la Matemática.* Madrid: Síntesis. Recuperable en:
`http://www.ugr.es/local/jgodino/Teoria_Metodos/haciateo.htm`

Godino, J. D. y Batanero, C. (1996). The dialectic relationships between research and practice: A meta-analysis of three research works. En: N. Malara (Ed.), *An International View of Didactics of Mathematics as a Scientific Discipline*, (pp. 13-22). Universidad de Módena. Versión en español recuperable en:
`http://www.ugr.es/local/jgodino/Teoria_Metodos/dialectes.htm`

Sierspinska, A. y Lerman, S. (1996). Epistemologies of mathematics and of mathematics education. En: A. J. Bishop et al. (eds.), *International Handbook of Mathematics Education* (pp. 827-876). Dordrecht, HL: Kluwer, A. P.

Juan D. Godino
Departamento de Didáctica de la Matemática
Facultad de Ciencias de la Educación, Universidad de Granada
Campus de Cartuja, 18071 Granada
correo electrónico: `jgodino@ugr.es`
Página web: `http://www.ugr.es/local/jgodino/`