Centro de Investigación Avanzada en Educación
Universidad de Chile

Documento de Trabajo N°4 , Octubre 2012

Red causal de contenidos matemáticos en cuarto básico: qué dice la evidencia empírica

Roberto Araya, Johan Van der Molen

Red causal de contenidos matemáticos en cuarto básico: qué dice la evidencia empírica

Roberto Araya[1], Johan Van der Molen[1]

[1] Centro de Investigación Avanzada en Educación, Universidad de Chile, Periodista José Carrasco Tapia. 75, Santiago, Chile.

Resumen: *Un problema central en el diseño curricular y en la planificación de clases es conocer cómo ciertos contenidos impactan en otros. Usualmente esta red de influencias causales se construye en base a las opiniones de diferentes expertos. Algunos de ellos contribuyen explicitando la dependencia lógica de los contenidos y otros contribuyen con sus experiencias personales al enseñar esos contenidos. Sin embargo, es muy importante contrastar esas opiniones con el proceso de aprendizaje que efectivamente ocurre en el aula y construir redes causales en base a la evidencia empírica. En este trabajo se analizan los registros almacenados de más de medio millón de ejercicios en línea realizados semanalmente en el 2011 por 805 estudiantes en 23 cursos de cuarto básico de 13 escuelas vulnerables. Con algoritmos de minería de datos y machine learning se construye automáticamente redes causales de contenidos para cuarto básico. Para cada contenido seleccionado del currículo de matemáticas se buscan entre todos los contenidos sus desempeños históricos que mejor predicen el desempeño futuro en el contenido seleccionado. La red así construida, es a nuestro conocimiento, la primera red causal de contenidos de un currículo construida empíricamente. Los patrones causales encontrados pueden ayudar al rediseño curricular y a la planificación de clases.*

Palabras claves: diseño curricular, planificación de clases, red de impactos causales entre contenidos, minería de datos educacionales, machine learning

Introducción

Si bien los contenidos matemáticos a enseñar tienen una estructura lógica bien definida en la que unos contenidos son prerrequisitos para otros o tienen una clara influencia en la posibilidad de extender y complementar otros, el aprendizaje es un fenómeno psicológico que no sigue estrictamente esa misma dependencia lógica. Existen conceptos que se olvidan y lo hacen con tasas de olvido que dependen de la estrategia que los estudiantes siguen para estudiar y ejercitar (Roediger & Karpicke, 2006; Karpicke & Blunt, 2011), hay preconcepciones erradas muy difíciles de superar (Geary, 2007), hay multiplicidad de concepciones y estrategias contradictorias que compiten en cada instante que se aborda un problema (National Math Panel, 2008), existen conceptos y estrategias aprendidas no conscientemente y otras que son conscientes en unos momentos pero no en otros (Siegler & Stern, 1998; Hadamard, 1945), hay errores sistemáticos muy difíciles de identificar, etc. Por ejemplo, solo en la resta existen documentados más de 100 errores (Van Lehn, 1990) que los estudiantes cometen y que una vez mecanizados cuesta enormemente corregir. No basta con señalar el error ni hacer ver la forma correcta de operar. Siegler & Araya (2005) da cuenta de varios mecanismos en el uso y descubrimiento de estrategias en aritmética. Existe inercia de estrategias anteriores, interrupciones de estrategias en curso, inhibiciones de unas estrategias sobre otras, influencias de percepciones sensoriales, limitaciones de memorias de corto plazo y de acceso a memorias de largo plazo.

Esta compleja dinámica está muy lejos de comprenderse bien, pero claramente no es un proceso lógico con un aprendizaje monótonamente creciente. Es un proceso reiterativo en que se avanza en forma de olas superpuestas (Siegler, 1998) con constantes regresiones a conceptos y estrategias supuestamente superadas, y con enormes influencias contextuales, sociales y muchas otras desconocidas o muy difíciles de controlar. Analogías y metáforas (Araya, et Al, 2010; Richland et Al, 2007) pueden apoyar la comprensión y el aprendizaje, pero también metáforas superficiales (Richland et Al, 2004) pueden confundir y obstaculizar el aprendizaje. El uso de representaciones visuales o pictóricas puede también tener una gran influencia (Hadamard, 1945; National Math Panel, 2008). Existen formas de presentación de conceptos que pueden hacer una gran diferencia, aún cuando la estructura lógica y matemática es completamente similar. Por ejemplo, en problemas de probabilidades, los formatos ecológicamente válidos (Gigerenzer, 2000) de frecuencias naturales logran una comprensión y tasas de errores mucho mejores a los que logran formatos simbólicos con fracciones o porcentajes. Toda esta red de dependencias y la dinámica asociada está muy lejos de la estructura lógica de los contenidos.

Adicionalmente al proceso psicológico de aprendizaje está el contexto social (The National Academies, 2001), en el cual la influencia del profesor y la de otros estudiantes es muy importante. El aprendizaje depende críticamente de cómo el profesor introduce los contenidos, los ejemplos y representaciones que utiliza, la frecuencia e intensidad que hace ejercitar a los estudiantes (Araya & Van der Molen, 2012), si hace que colaboren y resuelvan problemas en equipo. Por esta razón, las mediciones de los procesos psicológicos de aprendizaje realizados en laboratorios universitarios o de investigación tampoco son completamente suficientes. Es necesario incluir mediciones en aula en clases reales (U.S. Department of Education, 2010; The

National Academies, 2001), tal como se implementan los currículos, con las limitaciones y ventajas que poseen los profesores al estar frente a cursos reales en el ambiente escolar. Esto significa que en la implementación del currículo no se puede suponer que es un proceso independiente de la profesora, pues ella tiene una gran influencia. Sin embargo, es muy importante aislar los efectos particulares de los profesores de los efectos del currículo. Es decir, determinar cómo el currículo adaptado de diversas maneras por un gran conjunto de profesores termina impactando el aprendizaje y aún así determinar cómo los diferentes componentes del currículo influyen en los otros. Esto requiere considerar una muestra grande de diferentes cursos, cada uno con sus profesores que han implementado el currículo según sus propias experiencias, recursos, materiales y capacitaciones recibidas. En este trabajo analizamos la información de 23 cursos diferentes, cada uno con su propia profesora. Si bien todavía es un número pequeño de cursos y docentes, ya puede obtenerse patrones interesantes y estadísticamente significativos.

Es por lo tanto muy importante saber cuál es realmente la estructura de impactos de unos contenidos sobre otros, pero no basado en consideraciones teóricas o lógicas sino con la información tal como sucede en el proceso de aprendizaje con los estudiantes en las escuelas. Esto significa tomar datos empíricos de estudiantes en proceso de aprendizaje y buscar patrones de dependencia. Es importante señalar que el desafío no es exclusivamente matemático sino que tiene una fuerte componente de naturaleza muy distinta. Hay gran importancia de fenómenos psicológicos y sociales, y por tanto es un enorme desafío que debe enfrentarse principalmente de forma empírica.

Paradojalmente, la evidencia empírica está típicamente ausente en las decisiones de diseño curricular. Consecuentemente, en gran parte de la planificación de clases, que principalmente se realiza siguiendo las especificaciones de secuenciamiento de contenidos que viene descrito en el currículo, también está en gran medida ausente la evidencia empírica. Esta situación se explica en parte por la gran dificultad de conseguir la información detallada de los procesos de aprendizaje, tal como ocurre en las aulas. Sin embargo, avances recientes en la factibilidad de registrar información está comenzando a cambiar esta realidad (U. S. Department of Education, 2010). Existe todo un campo de investigación nueva, denominado Minería de Datos Educacionales, que precisamente se está haciendo cargo de explorar esta oportunidad. Se abre una gran posibilidad de lograr una aproximación más real y por tanto potencialmente más efectiva. La American Statistical Association (2007), por ejemplo, recientemente ha hecho un llamado a usar métodos científicos en la investigación en educación y el Departamento de Educación de EEUU (U.S. Department of Education, 2010) destaca como una tecnología clave la minería de datos educacionales, y establece que entre los principales objetivos de la minería de datos educacionales está el “descubrir o mejorar modelos que caractericen el contenido a ser aprendido y las secuencias instruccionales óptimas” (U.S. Department of Education, 2012b).

Cómo medir el impacto de un currículo

El impacto de un currículo depende críticamente de su implementación efectiva. Puede fácilmente suceder que un currículo teóricamente bien fundado falle debido a que no se implemente adecuadamente. La implementación requiere asegurar una capacitación completa de docentes y el acompañamiento y seguimiento hasta asegurar un grado de implementación razonable. Una forma relativamente barata de medir la implementación es mediante el seguimiento del uso de textos. Normalmente los textos han sido diseñados siguiendo el currículo y una comisión independiente de expertos revisa que el ajuste sea apropiado. Con encuestas y entrevistas a los docentes puede lograrse una primera estimación del uso de textos e indirectamente de la implementación del currículo. Una estrategia más realista pero más cara, es el registro de visitas a aulas, o mejor aún, el registro y análisis de videos de clases. Por ejemplo, en Araya y Dartnell (2008) se revisaron más de 700 videos enviados por docentes para la evaluación docente del año 2005. En cada video se examinó una muestra de 4 minutos que fue codificada por 4 evaluadores independientes. En este estudio se encontró que prácticamente ningún docente usaba los textos entregados gratuitamente por el Ministerio de Educación. En realidad la adopción de un material instruccional requiere muchas horas de capacitación. Según un estudio en EEUU (Banilower et Al, 2006) realizado para la National Science Foundation con encuestas a 75.000 docentes, 1.783 entrevistas a docentes y 1.620 observaciones a aulas realizadas en un espacio de 4 años, se encontró que para lograr un 50% de probabilidad de que un docente utilice en todas las clases un material instruccional se requiere al menos 200 horas de capacitación al año en ese material. Es decir, es más de un mes dedicado exclusivamente a capacitación en el uso específico de ese material. Por estas dificultades de adopción, es complejo medir el impacto del currículo y sus componentes a través de los textos y su uso.

Otra forma de ver el impacto de los diferentes componentes del currículo es a través de las pruebas nacionales como el SIMCE. Por ejemplo, Bhatt & Koedel (2011) comparan datos de 17 años de tres currículo implementados en Indiana. Estos son principalmente textos, con tecnologías complementarias y estrategias pedagógicas. Los autores pueden concluir las existencias de diferencias de impacto entre los tres currículos en 716 escuelas de Indiana. En general este tipo de pruebas estandarizadas nacionales o estatales son cuidadosamente diseñadas por comités de expertos, asegurándose de contener el currículo, y luego piloteadas en muestras grandes para verificar la adecuación como instrumento válido de medición. Sin embargo, estas pruebas se hacen una vez cada cuatro años por estudiante (en el caso del SIMCE), siendo la primera al final del cuarto básico. Son en total cerca de 35 preguntas, todas medidas en un mismo día y al final de año. Con esta información es prácticamente imposible determinar la red de causalidad de unos componentes del currículo sobre los otros. No basta con los resultados de los ejercicios, es importante registrar los momentos del año en que van ocurriendo los eventos de aprendizaje y evaluación.

El impacto de un currículo puede medirse mejor si la implementación del currículo está bien controlada. Esto permite diferenciar los efectos en el aprendizaje de los estudiantes de otros efectos. Un extenso estudio comparando el efecto de los tres programas de mejoramiento

educacional más ampliamente diseminados en EEUU (Consortium for Policy Research in Education, 2009) proveídos por tres organizaciones asesoras externas, en el que se analizó luego de cuatro años de aplicación su efecto en una muestra de 115 escuelas básicas con un total de 300 docentes, 800 directivos y 7.500 estudiantes, mostró que programas que enfatizan en la capacitación, el empoderamiento y autonomía del docente logran más satisfacción en los docentes y altos niveles de motivación de los docentes para mejorar, pero no logran mejorar el desempeño de los estudiantes. El nulo efecto en los desempeños de los estudiantes de este programa contrasta con el de los otros dos programas estudiados en los que el docente debe seguir una secuencia prediseñada y especificada en forma muy detallada, pero que le otorga muy poca autonomía al docente. En este último tipo de programas se observa que los estudiantes sí logran mejorías de desempeño. Este estudio nos alerta de que la forma de implementación de un currículo puede tener un enorme impacto. El estudio, sin embargo, sólo reporta el impacto global en el desempeño de los estudiantes. No establece para cada uno de las tres implementaciones, cómo cada componente del currículo influye en los otros componentes.

La red causal de aprendizajes entre contenidos del currículo

Para mejorar el diseño curricular es necesario lograr conocer no solamente el impacto global del currículo, sino que además conocer en mayor nivel de detalle la interacción e impacto en el aprendizaje entre los componentes del currículo. El desafío es medir cómo el aprendizaje en cada contenido depende del aprendizaje previo de componentes de otros contenidos del currículo. Según the American Statistical Association (2007), "entender los efectos causales de los regímenes instruccionales es central para la investigación aplicada en educación". Esto requiere considerar cuidadosamente la causalidad, y no solamente la correlación entre el aprendizaje de conceptos. No basta solamente saber que los desempeños de los estudiantes en dos conceptos son generalmente similares, ambos están bien o ambos están mal. Es necesario intentar averiguar cuál desempeño en un contenido causa o facilita cuál el desempeño en otro contenido.

Esto significa que en cada instante es necesario para cada contenido y su desempeño futuro encontrar cuál de los otros contenidos (o eventualmente él mismo) y sus desempeños pasados influyen en ese contenido futuro. Por ejemplo, en el mes de Junio, cuál de los contenidos y sus desempeños correspondientes en los meses anteriores a Junio permite predecir mejor el comportamiento de fracciones desde Junio a Agosto. Luego ver si el patrón encontrado se repite en Julio en lugar de Junio, y así en varios otros meses.

En el análisis que sigue se utilizan los contenidos mínimos obligatorios (CMOs) definidos oficialmente en el currículo para el año 2011 por el Ministerio de Educación. En base a los registros de los estudiantes al resolver en un sistema en línea cientos de ejercicios, se busca determinar una red de los impactos que causan unos contenidos sobre otros, es decir cómo el desempeño anterior de uno predice el desempeño futuro de otro. Esta construcción depende de la base de ejercicios realizados y de la clasificación de cada ejercicio, donde manualmente a cada ejercicio se le asignó uno o más contenidos. Depende también de la secuencia de ejercicios que los profesores escogieron y de los estudiantes que realizaron los ejercicios durante el año. Aún así, dado que son

23 profesores, cada uno actuando en forma relativamente independiente de los otros, y dado que son 805 estudiantes, es esperable que la red construida debería ser razonablemente estable. Es decir en pares meses-estudiantes que no se utilizaron para la búsqueda de patrones de causalidad entre contenidos es necesario verificar si se mantienen esos mismos patrones. Un testeo más exigente, sería ver si esos mismos patrones se mantienen con otros estudiantes, otros profesores y otras escuelas.

En las predicciones existen dos tipos de errores. Por una parte están desempeños futuros buenos en los próximos 3 meses que se predicen como malos y, por otra parte están comportamientos futuros malos que se predicen como buenos. En lo que sigue para cada par alumno-mes definimos como desempeño bueno en pruebas en una categoría de contenidos si en los próximos tres meses el alumno sacó un puntaje sobre 50% en los ejercicios que contenían esa categoría de contenidos. En tareas consideramos que un alumno-mes tuvo un desempeño bueno si más de la mitad de las veces que le tocó ejercicios que contenían la categoría de contenido los hizo correctamente al primer intento.

Para una variable histórica v y en un punto de corte c de un contenido B (por ejemplo, errores al hacer ejercicios de B) como medida de la capacidad de v de discriminar el buen del mal desempeño en un contenido A consideraremos el módulo de la diferencia entre el porcentaje de desempeños buenos en A que cumplen v >c (es decir, podríamos interpretarlos como aciertos en predicción con v y c dentro de los buenos desempeños futuros) y el porcentaje de desempeños malos que cumplen v > c (que podríamos interpretarlos como errores en predicción con v y c dentro de los malos desempeños futuros). Es decir, tomamos el porcentaje de aciertos con los buenos desempeños futuros y le descontamos el porcentaje de desacierto con los malos desempeños futuros. En la figura 1, se ilustra esta métrica.

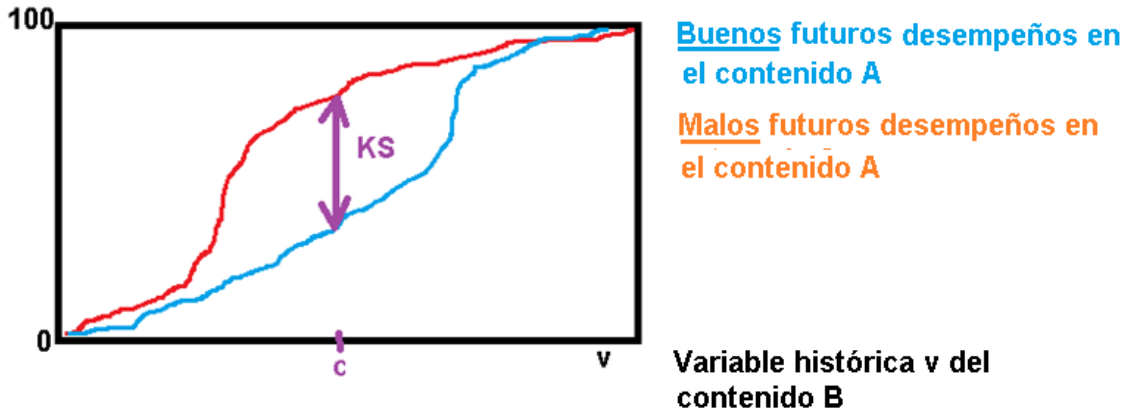


Figura 1: Métrica KS para variable histórica v *con punto de corte* c*. Para los valores de* v *menores a* c *se acumulan muchos más malos desempeños que buenos desempeños futuros. El módulo de la diferencia es el KS.*

Este es un número entre 0 y 100. En caso de ser 100 significa que no hay error en la predicción, todos los buenos se predicen como buenos, y todos los malos se predicen como malos. Por el contrario, si esta medida es cero, entonces significa que la capacidad de predicción es nula. Por ejemplo, la mitad de los buenos se predice como bueno (o sea 50% de los buenos se predice como bueno) y la mitad de los malos se predice como buenos (50% de los malos se predice como bueno). Una propiedad importante de esta medida es que el resultado es el mismo si consideramos el módulo de la diferencia entre el porcentaje de comportamientos malos que cumplen v <c y el porcentaje de comportamientos buenos que cumplen con v <c. Luego buscamos el punto de corte c donde se obtiene el mejor KS y ese valor de KS queda asociado a la variable v. Esta medida de discriminación es conocida como la distancia Kolmogorov- Smirnov (Friedman, 1977). Es importante observar que hay muchas otras medidas de discriminación usada (Gini, purezas, etc). Sin embargo, estudios comparativos empíricos (Mingers, 1989) muestran que las diferentes métricas no producen resultados muy diferentes.

Es importante observar que esta medida tiene incorporada la causalidad: cómo lo anterior predice lo que viene, y no es sólo una medida de correlación. Aquí la ligazón causa-efecto está marcada por la temporalidad. Si bien ocasionalmente puede darse por azar que un buen (mal) desempeño en un contenido preceda un buen (mal) desempeño en otro, es muy poco frecuente que esta ocurrencia se produzca reiteradamente por azar en diversos meses de año y con cientos de estudiantes de escuelas distintas. Si así es, entonces quiere decir que cierto grado de causalidad y no es mera coincidencia producto del azar. De todas formas, es muy importante asegurar la estabilidad o robustez de los hallazgos. Esto se logra separando una muestra de datos de desempeños de estudiantes en ciertos meses que no sean usados para construir la red. Esta muestra es la que se denomina base de datos de validación. Ella luego se utiliza para verificar si también se cumplen los mismos patrones ya encontrados en la base de construcción y si las métricas de causalidad son similares.

Fuente de información

Consideramos los registros de 515.823 ejercicios realizados por 805 estudiantes de 23 cursos de cuarto grado de educación básica. Los ejercicios fueron realizados durante todo el año 2011, desde Marzo a Diciembre. Los estudiantes asistían a un par de sesiones semanales a un laboratorio tic donde realizaban los ejercicios. En total son 13 escuelas con gran cantidad de alumnos vulnerables, pertenecientes a tres comunas de nivel socio económico bajo y medio bajo de Santiago. En total comprenden 23 docentes, cada uno a cargo de un curso. De los 515.823 ejercicios, 353.575 fueron realizados en modo tarea, es decir con retroalimentación inmediata (en menos de un segundo) que el sistema le entrega al estudiante mientras hace el ejercicio. El resto de los ejercicios fueron hechos en modo prueba, en los que el estudiante no tiene retroalimentación inmediata si no posiblemente al terminar la sesión o en otra sesión.

Los estudiantes realizaron los ejercicios en un laboratorio computacional que cuenta con el apoyo de una docente que está específicamente a cargo de operar la plataforma de ejercicios en línea. Por otra parte, el docente a cargo del curso trae los estudiantes al laboratorio y se dedica

solamente a apoyar a los estudiantes con consultas de contenidos. Se utilizó una plataforma de ejercitación en línea que contiene un catálogo de cerca de 2.000 ejercicios de matemáticas para primer ciclo de enseñanza básica. Los ejercicios son de diferente naturaleza y no sólo de selección múltiple. Por ejemplo, algunos ejercicios solicitan ubicar un número en la recta numérica habiendo infinitas posiciones posibles donde hacer clic. Otros ejercicios requieren ingresar números que son resultados de cálculos o digitar palabras. Otros requieren revisar una lista de pasos o argumentos y hacer clic en el que está errado o hacer clic indicando que está todo bien. Existen otros ejercicios que requieren hacer múltiples clics describiendo una secuencia ordenada de posiciones en la pantalla, como por ejemplo las que se requiere para hacer un gráfico. Tal como lo recomienda (U.S. Department of Education, 2012a), hay ejercicios rutinarios que pretenden ayudar a memorizar (como las tablas de multiplicar) y también hay ejercicios no rutinarios. Existen ejercicios que pretenden incentivar el uso de metáforas (Araya, 2000) como por ejemplo que la suma y resta son traslaciones, y en particular representaciones visuales y el tránsito de lo concreto, a lo pictórico y luego a lo simbólico. Otros ejercicios promueven el uso de formatos ecológicamente válidos (Gigerenzer, 2000; Araya 2000), como por ejemplo el uso de frecuencias naturales para estimar probabilidades y el uso de intercambios y frecuencias temporales para comparar fracciones. Hay ejercicios tipo puzles, sudokus y laberintos. También hay ejercicios que promueven competencias de modelamiento matemático (OECD, 2003; Araya et Al, 2012), como por ejemplo detectar y describir patrones, etc. Aún así, si algún docente estima que faltan ejercicios de los que él cree importante, entonces puede entregar esos ejercicios escritos en un procesador de texto y luego durante la semana el equipo de soporte ingresa una versión de los ejercicios ajustados al formato del sistema.

Cada ejercicio está catalogado según eje y puede tener asociado uno o más contenidos mínimos obligatorios (CMO), según la lista de CMOs especificados en el ministerio de Educación para 2010 y 2011. Cada semana el docente del laboratorio coordina con el docente del curso el tipo de ejercicios a asignar a los alumnos. La asignación la hace el docente de laboratorio desde su casa o en el mismo laboratorio en la escuela.

Durante la sesión el docente monitorea en un tablet o desde un netbook el progreso de los estudiantes. En cada instante un sistema de alerta lista los estudiantes que están teniendo más dificultades ordenados por el tiempo que llevan sin avanzar y el número de intentos infructuosos que están teniendo. Junto a cada estudiante se agrega el ejercicio que lo tiene detenido y el tiempo que lleva en él. De esta manera el docente sabe qué estudiante requiere atención personalizada y en qué ejercicio específico, lo que queda registrado para futuras estadísticas. El sistema de alerta también lista los ejercicios que más complican a toda la clase, permitiéndole saber si hay uno muy difícil, con problemas de redacción u otro problema y que entonces amerite que el docente intervenga y aclare los conceptos o el problema a toda la clase. Todas las actividades de la sesión forman parte de un reporte diario que se envía automáticamente por email al docente y jefe de la Unidad Técnica Pedagógica (UTP) al fin del día. La plataforma está diseñada para conducir el avance de la clase completa como un todo y no solamente dejar a que los estudiantes avancen solos. Tiene facilidades para promover la cooperación y que estudiantes adelantados apoyen a sus pares con más dificultades, los que son detectados por el sistema de

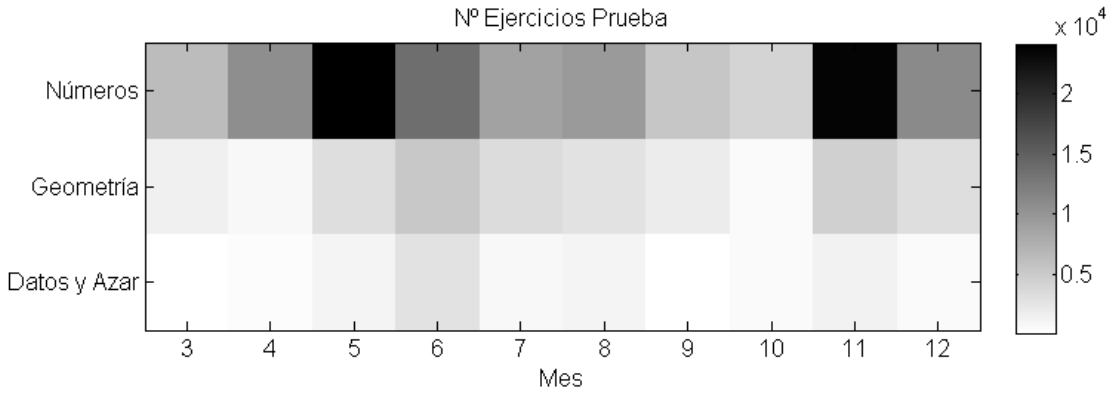
alerta. Una vez que un estudiante ha terminado correctamente sus ejercicios la docente puede asignarlo como monitor y el sistema le asigna pares a ayudar.

Para cada ejercicio con retroalimentación inmediata está almacenado el número de intentos hasta realizarlo correctamente y el tiempo que le tomó terminarlo completamente a cada estudiante. En el caso de ejercicios sin retroalimentación (tipo prueba) se tiene almacenado el porcentaje de respuestas correctas del ítem y su tiempo de respuesta. Los ejercicios con tiempos de respuesta de menos de 3 segundos o que estén a menos de dos desviaciones estándares de la media fueron sacados del análisis. Estos corresponden en general a ejercicios que se abrieron y cerraron muy rápidamente y por lo tanto no se realizaron. Además, los contenidos que tuvieran menos de 200 pares de alumnos-meses asociados en la base de construcción o menos de 200 alumnos-meses en la base de validación no son considerados en el análisis.

Procedimiento

La gran cantidad de información que se registra en la medida que los estudiantes resuelven ejercicios en el sistema en línea es una excelente oportunidad para buscar patrones y construir modelos predictivos, lo que se denomina minería de datos educacionales (U.S. Department of Education, 2012b). En este trabajo reportamos la búsqueda de patrones realizada automáticamente por un algoritmo que contiene métodos de machine learning tipo inducción de árboles de decisión (Michie et Al, 1994; Araya & Gigon, 1992) y support vector machines (Bennet & Campbell, 2000).

Los gráficos siguientes muestran el número de ejercicios realizados cada mes. La Figura 2 muestra en el gráfico de arriba el número de ejercicios tipo prueba para cada uno de los tres ejes. En el gráfico de abajo se muestra el número de ejercicios tipo tarea para cada uno de los tres ejes.



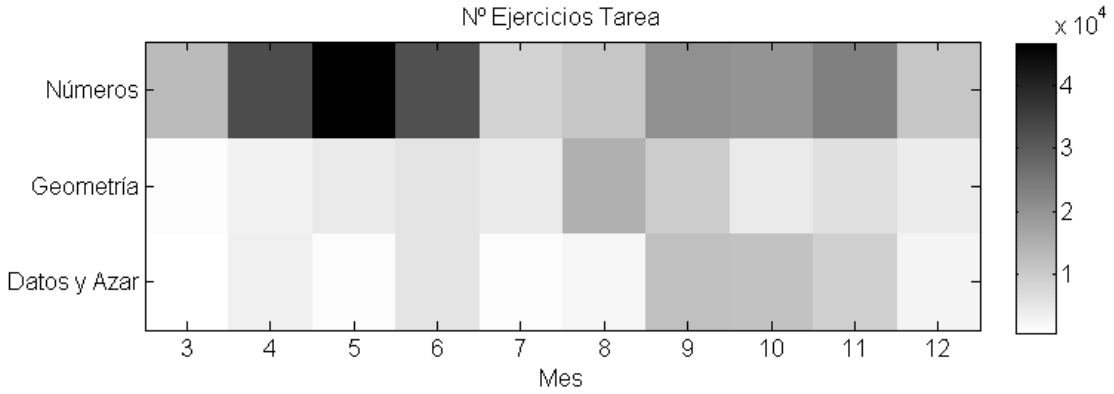


Figura 2: Número de ejercicios hechos de los tres ejes por cada mes. Arriba sin los ejercicios tipo prueba y abajo se muestran los ejercicios tipo tarea.

Se observa que se realizaron más ejercicios tanto del tipo pruebas como de tareas en el eje de números, y que en el mes de Mayo y Noviembre hubo mayor número de ejercicios tipo pruebas de ese eje. En tareas se observa un gran número de ejercicios el primer semestre. Se ve también que en Agosto sube el número de ejercicios de geometría y hacia fin de año sube el número de ejercicios en el eje azar.

La figura 3 muestra el número de ejercicios del tipo tarea realizados en cada mes por cada uno de los 17 contenidos con más actividad. Se observa, por ejemplo, una gran cantidad de ejercicios en cálculo mental y escrito con números naturales en el primer semestre, y tablas, pictogramas y gráficos en el segundo semestre.

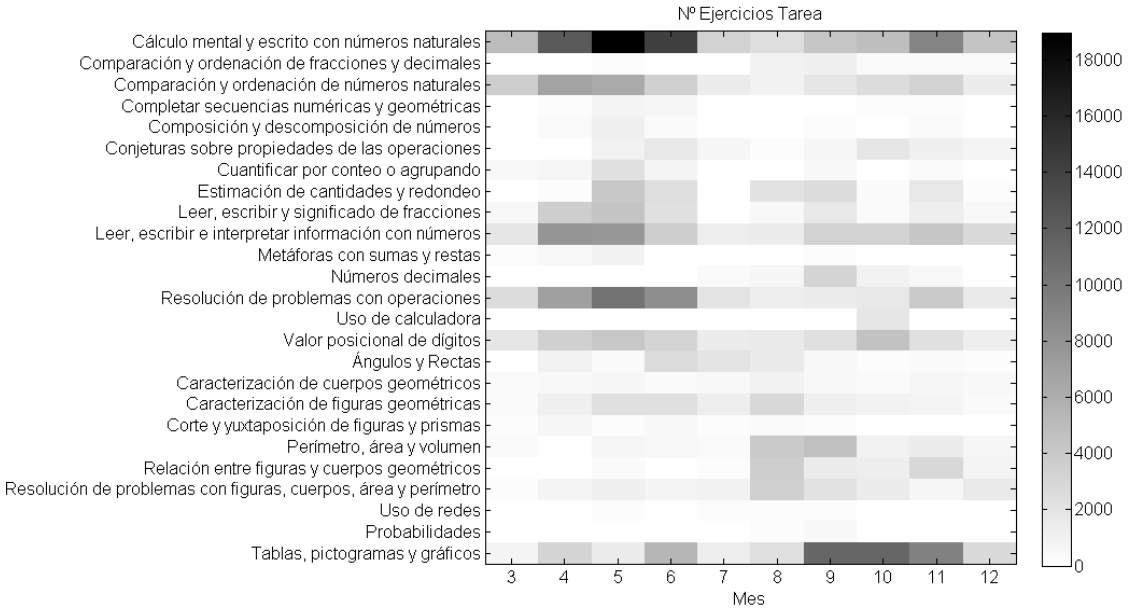


Figura 3: Número de ejercicios hechos por cada mes para los contenidos con más registros de actividad.

El algoritmo principal de machine learning usado es uno de inducción de árboles de decisión, en el que se busca la variable que mejor predice el futuro y segmenta la base de datos de acuerdo a esa variable. Más específicamente, el algoritmo revisa para cada contenido (Tabla 1) cuáles contenidos históricos y cuáles de todas las variables asociadas a esos contenidos (Tabla 2) predice mejor el contenido futuro en cuestión. Para realizar esto, para cada estudiante y mes a analizar se registra el comportamiento futuro del estudiante en cada contenido. La Tabla 1 muestra la lista de todos los 75 contenidos (CMOs) oficiales del currículo de primero a cuarto básico considerados y su agrupación en 37 categorías que se utilizan en este trabajo.

Contenido (CMO) oficial del currículo	Nivel	Categoría	Eje
Cálculo mental: combinaciones aditivas básicas, su extensión a números de 2 cifras y aplicación a situaciones significativas.	Primero Básico	Cálculo mental y escrito con números naturales	Números
Cálculo mental: combinaciones aditivas con números de 2 y 3 cifras, estrategias de cálculo basadas en descomposiciones aditivas y en las propiedades de las operaciones, aplicación a situaciones significativas.	Segundo Básico	Cálculo mental y escrito con números naturales	Números
Cálculo mental y escrito de adiciones y sustracciones con números naturales en el nuevo ámbito estudiado extensión de las combinaciones aditivas básicas, generalización de las estrategias de cálculo mental y de los algoritmos de cálculo escrito.	Tercero Básico	Cálculo mental y escrito con números naturales	Números
Cálculo mental de productos y cuocientes de números naturales con resultados en el conjunto de los números naturales: combinaciones multiplicativas básicas, multiplicaciones por potencias de 10, aplicación a situaciones problemáticas.	Tercero Básico	Cálculo mental y escrito con números naturales	Números
Cálculo escrito de productos en que uno de los factores es un número natural de una o dos cifras o múltiplo de 10, 100 o 1.000, uso de un algoritmo de cálculo basado en la descomposición aditiva de los factores y la propiedad distributiva de la multiplicación sobre la adición.	Tercero Básico	Cálculo mental y escrito con números naturales	Números
Cálculo escrito de cuocientes y restos en divisiones donde el divisor es un número de una cifra: uso de un algoritmo de cálculo basado en el carácter inverso de la división con respecto a la multiplicación.	Tercero Básico	Cálculo mental y escrito con números naturales	Números
Cálculo mental y escrito en situaciones donde sea significativo realizar este tipo de cálculo, de operaciones combinadas con números naturales en el ámbito numérico estudiado: extensión de las estrategias de cálculo mental conocidas y de los algoritmos de cálculo escrito.	Cuarto Básico	Cálculo mental y escrito con números naturales	Números
Relación de los términos “mayor”, “menor” o “igual que”, con la adición y sustracción. Ordenación de los números naturales estudiados y comparación de cantidades utilizando dichos términos.	Primero Básico	Comparación y ordenación de números naturales	Números
Comparación de cantidades y ordenamiento de los números naturales estudiados utilizando los términos “igual que”, “mayor que” y “menor que”, describiendo la estrategia utilizada.	Segundo Básico	Comparación y ordenación de números naturales	Números
Representación de números naturales en la recta numérica y empleo de los símbolos <, > e = para ordenar y comparar números, cantidades y medidas, dentro del ámbito numérico estudiado.	Tercero Básico	Comparación y ordenación de números naturales	Números
Representación de números naturales o subconjuntos de ellos en la recta numérica y empleo de los símbolos <, > e = para ordenar y comparar números naturales dentro del ámbito numérico estudiado cantidades y medidas.	Cuarto Básico	Comparación y ordenación de números naturales	Números
Identificación de los valores desconocidos en secuencias numéricas, dadas las reglas que las generan.	Primero Básico	Completar secuencias numéricas y geométricas	Números
Continuación de patrones numéricos y geométricos, dada una regla que lo genera, y determinación de valores desconocidos en secuencias numéricas.	Segundo Básico	Completar secuencias numéricas y geométricas	Números
Identificación de reglas que generan secuencias numéricas o geométricas, cuyos elementos están ordenados de acuerdo con un patrón determinado, y continuación de estas secuencias según la regla.	Tercero Básico	Completar secuencias numéricas y geométricas	Números
Formulación y verificación de conjeturas, en casos particulares, acerca de la posibilidad de que una secuencia sea generada por más de una regla o patrón.	Cuarto Básico	Completar secuencias numéricas y geométricas	Números
Elaboración de estrategias basadas en la composición y descomposición aditiva para el cálculo escrito de adiciones y sustracciones.	Primero Básico	Composición y descomposición de números	Números
Establecimiento de estrategias basadas en la descomposición aditiva y en las propiedades de las operaciones para el cálculo escrito de adiciones y sustracciones.	Segundo Básico	Composición y descomposición de números	Números
Extensión de la composición y descomposición aditiva a números del nuevo ámbito.	Tercero Básico	Composición y descomposición de números	Números
Formulación y verificación de conjeturas respecto a: relación inversa de la sustracción respecto de la adición y viceversa, conmutatividad y asociatividad de la adición, comportamiento del 0 (cero) en adiciones y sustracciones.	Segundo Básico	Conjeturas sobre propiedades de las operaciones	Números

Formulación y verificación de conjeturas, en casos particulares, acerca de la relación entre la multiplicación y la adición; representación a través de la operación de multiplicación de situaciones correspondientes a aportes equitativos, arreglos rectangulares (elementos ordenados en filas y columnas); situaciones de correspondencia uno a varios y su aplicación en la resolución de problemas.	Tercero Básico	Conjeturas sobre propiedades de las operaciones	Números
Formulación y verificación de conjeturas, en casos particulares, respecto a la relación entre la división y la sustracción; representación, a través de la operación de división, de situaciones correspondientes a repartos equitativos; comparación por cuociente y su aplicación en la resolución de problemas.	Tercero Básico	Conjeturas sobre propiedades de las operaciones	Números
Formulación y verificación de conjeturas, en casos particulares, relativas a: relación entre la multiplicación y la división, propiedad conmutativa y asociativa de la multiplicación, propiedad distributiva de la multiplicación respecto a la adición, comportamiento del 0 y el 1.	Cuarto Básico	Conjeturas sobre propiedades de las operaciones	Números
Establecimiento de estrategias para cuantificar elementos del entorno mediante el conteo de cantidades de 1 en 1 o por agrupaciones.	Primero Básico	Cuantificar por conteo o agrupando	Números
Establecimiento de estrategias para la resolución de problemas referidos al conteo de cantidades por agrupaciones (de 10 en 10, 15 en 15, 20 en 20, etc.).	Segundo Básico	Cuantificar por conteo o agrupando	Números
Determinación de valores desconocidos en igualdades de expresiones aditivas dentro del ámbito numérico conocido.	Segundo Básico	Determinar valores desconocidos en operaciones	Números
Determinación de valores desconocidos en igualdades de expresiones multiplicativas dentro del ámbito numérico conocido.	Tercero Básico	Determinar valores desconocidos en operaciones	Números
Estimación de una cantidad o medida a partir de referentes dados y aplicación a situaciones problemáticas en contextos cercanos.	Primero Básico	Estimación de cantidades y redondeo	Números
Estimación de una cantidad a partir de referentes dados y aplicación a situaciones problemáticas en contextos cercanos.	Segundo Básico	Estimación de cantidades y redondeo	Números
Redondeo de números y su aplicación para estimar cantidades, medidas, el resultado de operaciones o para detectar eventuales errores de cálculo.	Tercero Básico	Estimación de cantidades y redondeo	Números
Redondeo de números y su aplicación para estimar cantidades o medidas, el resultado de operaciones o para detectar eventuales errores de cálculo.	Cuarto Básico	Estimación de cantidades y redondeo	Números
Reconocimiento, lectura y escritura de números del 0 al 100	Primero Básico	Leer, escribir e interpretar información con números	Números
Interpretación de información expresada con números del 0 al 100 en contextos familiares y uso de estos números para comunicar información.	Primero Básico	Leer, escribir e interpretar información con números	Números
Reconocimiento, lectura y escritura de números naturales del 0 al 1.000; e, identificación de regularidades que se presentan en los nombres y escritura de esos números.	Segundo Básico	Leer, escribir e interpretar información con números	Números
Interpretación de información expresada con números del 0 al 1.000 en contextos familiares y uso de estos números para comunicar información.	Segundo Básico	Leer, escribir e interpretar información con números	Números
Lectura, escritura y formación de números naturales hasta el 100.000, generalizando los conocimientos adquiridos con relación a los números de 1, 2 y 3 cifras, interpretar información expresada con estos números y su uso para comunicar información.	Tercero Básico	Leer, escribir e interpretar información con números	Números
Lectura, escritura y formación de números naturales hasta el 1.000.000 a partir de los conocimientos adquiridos; interpretación de información proporcionada a través de dichos números y su empleo para comunicar información en diversos contextos.	Cuarto Básico	Leer, escribir e interpretar información con números	Números
Significado, lectura y escritura de fracciones simples o de uso frecuente (1/2,1/3, 1/4, 1/8, 3/4, 1/10, 1/100), su empleo para cuantificar y comparar partes de un objeto, de una unidad de medida o de una colección de elementos en contextos cotidianos; comparación entre fracciones y representación en la recta numérica.	Cuarto Básico	Leer, escribir y significado de fracciones	Números
Asociación de situaciones que implican juntar, agregar, avanzar, separar, quitar y retroceder con las operaciones de adición y sustracción.	Primero Básico	Metáforas con sumas y restas	Números
Lectura, escritura y reconocimiento del valor representado por cada dígito en números decimales entre 0 y 1 (hasta las cifras de las centésimas) y su relación con fracciones (1/2, 1/4, 3/4, 1/10, 1/100), empleo para cuantificar magnitudes, comparación entre números decimales y representación en la recta numérica.	Cuarto Básico	Números decimales	Números
Resolución de problemas referidos a contextos familiares y significativos con datos explícitos que contribuyan al conocimiento de sí mismos y del entorno, enfatizando en habilidades que dicen relación con la comprensión de la situación problemática, la selección y aplicación de la operación a utilizar y la identificación del resultado como solución al problema planteado.	Primero Básico	Resolución de problemas con operaciones	Números
Resolución de problemas en contextos familiares, con datos explícitos que contribuyan al conocimiento de sí mismos y del entorno, enfatizando en habilidades que dicen relación con la comprensión de la situación problemática, la selección y aplicación de la operación a utilizar para su solución y la identificación del resultado como solución al problema planteado.	Segundo Básico	Resolución de problemas con operaciones	Números
Resolución de problemas aplicando la estimación, comparación de cantidades y operaciones conocidas en los números naturales, en contextos cotidianos.	Tercero Básico	Resolución de problemas con operaciones	Números
Formulación y verificación de conjeturas, en casos particulares, y resolución de problemas en contextos significativos, haciendo uso de la estimación y comparación de cantidades y medidas y de las operaciones conocidas en el ámbito de los números naturales hasta el 1.000.000.	Cuarto Básico	Resolución de problemas con operaciones	Números

Uso de calculadora u otras herramientas tecnológicas para el estudio de regularidades numéricas y para facilitar el cálculo, utilizando como criterios el número de operaciones, el tamaño de los números y la complejidad de los cálculos.	Tercero Básico	Uso de calculadora	Números
Uso de calculadora u otras herramientas tecnológicas para el estudio de regularidades numéricas y para facilitar el cálculo numérico utilizando como criterios la cantidad de cálculos a realizar, el tamaño de los números y la complejidad de los cálculos.	Cuarto Básico	Uso de calculadora	Números
Determinación del valor representado por cada dígito en números naturales de dos cifras de acuerdo con su posición y su relación con los conceptos de unidad y decena.	Primero Básico	Valor posicional de dígitos	Números
Determinación del valor representado por cada dígito en números naturales de dos y tres cifras de acuerdo con su posición y su relación con los conceptos de unidad, decena y centena.	Segundo Básico	Valor posicional de dígitos	Números
Reconocimiento del valor representado por cada dígito en números hasta el 100.000, de acuerdo con su posición y su relación con los conceptos de unidad de mil, decena de mil y centena de mil.	Tercero Básico	Valor posicional de dígitos	Números
Reconocimiento del valor representado por cada dígito en números hasta el 1.000.000 de acuerdo con su posición y su relación con los conceptos de unidad de mil, decena de mil y centena de mil.	Cuarto Básico	Valor posicional de dígitos	Números
Identificación de ángulos menores, mayores e iguales al ángulo recto, así como también de rectas paralelas, perpendiculares y oblicuas.	Segundo Básico	Ángulos y Rectas	Geometría
Exploración de prismas rectos de base triangular o rectangular; identificación de sus caras, aristas y vértices; y, caracterización en función del número y forma de las caras.	Primero Básico	Caracterización de cuerpos geométricos	Geometría
Exploración de pirámides, cilindros y conos para su caracterización en función de las superficies y líneas que los delimitan.	Tercero Básico	Caracterización de cuerpos geométricos	Geometría
Reconocimiento de formas geométricas en el entorno y su descripción mediante un lenguaje geométrico básico; e, identificación de líneas rectas y curvas en estas formas.	Primero Básico	Caracterización de figuras geométricas	Geometría
Identificación de lados y vértices en polígonos y caracterización en función del número de lados.	Primero Básico	Caracterización de figuras geométricas	Geometría
Identificación y caracterización de cuadriláteros y triángulos en función del paralelismo, perpendicularidad y longitud de los lados. Formulación y verificación de conjeturas respecto a la relación entre longitud y paralelismo de lados en cuadriláteros.	Segundo Básico	Caracterización de figuras geométricas	Geometría
Formación y transformación de figuras planas mediante yuxtaposición y corte de formas triangulares y rectangulares; transformación de cuerpos geométricos mediante yuxtaposición y separación de prismas rectos.	Segundo Básico	Corte y yuxtaposición de figuras y prismas	Geometría
Estimación y medición de longitudes de objetos o distancias entre dos puntos, utilizando unidades de medida informales, tales como la medida de manos o pies o unidades estandarizadas como el metro, centímetro y milímetro e interpretación de información referida a longitudes.	Segundo Básico	Perímetro, área y volumen	Geometría
Interpretación de información referida a perímetros, en situaciones significativas, y determinación de la medida del perímetro en polígonos, expresando el resultado en metros, centímetros o milímetros.	Tercero Básico	Perímetro, área y volumen	Geometría
Interpretación de información relativa a áreas en contextos significativos y empleo de cuadrículas para cuantificar o estimar el área de rectángulos o de figuras que pueden descomponerse en rectángulos.	Cuarto Básico	Perímetro, área y volumen	Geometría
Representación en el plano de la elevación, perfil y planta de cuerpos geométricos, y recíprocamente trazado de la representación de dichos cuerpos geométricos en el plano a partir de sus vistas.	Cuarto Básico	Relación entre figuras y cuerpos geométricos	Geometría
Resolución de problemas que implican comparar características de figuras planas y prismas rectos de base triangular o rectangular.	Primero Básico	Resolución de problemas con figuras, cuerpos, área y perímetro	Geometría
Resolución de problemas que implican comparar características de triángulos y cuadriláteros, combinar y descomponer formas geométricas empleando cortes, dobleces o yuxtaposiciones; medición, adición, sustracción y estimación de longitudes.	Segundo Básico	Resolución de problemas con figuras, cuerpos, área y perímetro	Geometría
Formulación y verificación de conjeturas, en casos particulares, acerca de la posibilidad de armar cuerpos a partir de distintas redes y resolución de problemas referidos al cálculo de perímetros en situaciones significativas.	Tercero Básico	Resolución de problemas con figuras, cuerpos, área y perímetro	Geometría
Formulación y verificación de conjeturas, en casos particulares; resolución de problemas referidos a representaciones bidimensionales de cuerpos; estimación y cálculo de áreas utilizando cuadrículas.	Cuarto Básico	Resolución de problemas con figuras, cuerpos, área y perímetro	Geometría
Identificación y empleo de redes que permiten construir cuerpos geométricos.	Tercero Básico	Uso de redes	Geometría
Recolección de datos cuantitativos o cualitativos sobre objetos, personas y animales del entorno escolar y familiar; clasificación según uno o más atributos.	Primero Básico	Estadística	Datos y Azar
Representación de datos cuantitativos o cualitativos, en tablas y pictogramas simples, referidos a atributos de objetos, personas y animales del entorno escolar y familiar.	Primero Básico	Tablas, pictogramas y gráficos	Datos y Azar
Resolución de problemas en los cuales es necesario extraer información cuantitativa desde tablas y pictogramas simples construidos con datos provenientes desde el entorno escolar y familiar y comparación de conclusiones a partir de la información extraída desde diferentes tablas.	Primero Básico	Tablas, pictogramas y gráficos	Datos y Azar
Representación de datos cuantitativos o cualitativos, en tablas de doble entrada y pictogramas, recolectados sobre objetos, personas y animales del entorno escolar y familiar, y argumentación sobre la elección de las representaciones.	Segundo Básico	Tablas, pictogramas y gráficos	Datos y Azar
Resolución de problemas en los cuales es necesario extraer información desde tablas de doble entrada y pictogramas, que contienen datos cuantitativos extraídos desde el entorno escolar o familiar, para responder a preguntas planteadas.	Segundo Básico	Tablas, pictogramas y gráficos	Datos y Azar

Discusión sobre la utilidad de las tablas y gráficos para resumir y comunicar información referida a diversos temas y situaciones.	Segundo Básico	Tablas, pictogramas y gráficos	Datos y Azar
Representación de datos cuantitativos en tablas y gráficos de barras simples, recolectados desde el entorno social y cultural, e interpretación en forma verbal o escrita de dicha representación. Discusión sobre el tipo de información que se puede representar a través de tablas y gráficos de barras simples.	Tercero Básico	Tablas, pictogramas y gráficos	Datos y Azar
Resolución de problemas en los cuales es necesario extraer información desde tablas y gráficos de barras simples y formulación de afirmaciones respecto a los datos a los que hacen referencia.	Tercero Básico	Tablas, pictogramas y gráficos	Datos y Azar
Formulación de preguntas y propuestas de respuestas a situaciones de la realidad, mediante la observación de tablas y gráficos de barras simples, construidos con datos recolectados con relación a dichas situaciones.	Tercero Básico	Tablas, pictogramas y gráficos	Datos y Azar
Producción y comunicación de información a partir de datos organizados en tablas y gráficos de barras simples, tanto verticales como horizontales. Discusión sobre el tipo de datos que se puede representar a través de tablas y gráficos de barras simples.	Cuarto Básico	Tablas, pictogramas y gráficos	Datos y Azar
Resolución de problemas en los cuales es necesario extraer información desde tablas y gráficos de barras simples verticales y horizontales, comparación y formulación de afirmaciones respecto a las situaciones o fenómenos a los que se hace referencia.	Cuarto Básico	Tablas, pictogramas y gráficos	Datos y Azar

Tabla 1: Categorías de contenidos usados en el análisis y su relación con los 75 contenidos (CMOs) oficiales del Ministerio.

Para cada mes de análisis se calculan las diferentes variables de comportamiento histórico, y luego se calcula la métrica KS de todas las variables históricas anteriores al mes de análisis para todas las categorías de contenidos. Por ejemplo, para la categoría de contenido Completar secuencias numéricas y geométricas, se revisa para cada mes del año desde Mayo a Septiembre, el comportamiento futuro de cada estudiante y se determina si el desempeño fue bueno o malo. Un estudiante puede haber tenido un mal desempeño en Mayo pero uno bueno en Agosto, esto porque si bien en ese contenido en ejercicios sin retroalimentación (tipo prueba) le fue mal en Junio, Julio y Agosto, sin embargo en Septiembre, Octubre y Noviembre su desempeño fue bueno. De esta manera los 805 estudiantes en cada mes de Mayo hasta Septiembre quedan catalogados si tuvieron buen desempeño en los tres meses siguientes o si tuvieron mal desempeño. En total son 805*5= 4.025 pares alumno-mes clasificados para ejercicios tipo prueba y 4.025 pares alumno-mes clasificad clasificados para ejercicios tipo tarea, por cada uno de las 37 categorías de contenidos. Es decir, un total de 297.850 clasificaciones. Por otra parte, para cada contenido c y mes de análisis m se calculan 18 variables históricas, más la variable mes, indicadas en la Tabla 2.

Tipo	Variable	Descripción
Mes	mes	mes del año (mayo, junio, julio, agosto o septiembre) que delimita la historia y el futuro para cada observación. Se considera historia desde marzo hasta el valor de "mes", incluyéndolo.
Puntaje	tasa_ptje_p_c	(suma de puntos obtenidos en los ejercicios modo prueba asociados al cmo c hasta el mes m)/(suma de puntajes maximos en los ejercicios modo prueba asociados al cmo c hasta el mes m)
Puntaje	prom_ptje_p_c	promedio de la variable [puntaje/puntaje máximo] sobre todos los ejercicios modo prueba asociados al cmo c hasta el mes m
Puntaje	corr_p_c	(nº de ejercicios en el modo prueba asociados al cmo c en que el puntaje obtenido es igual al maximo hasta el mes m)/(total ejercicios en el modo prueba asociados al cmo c hasta el mes m)
Puntaje	pregsobre50_p_c	(nº de ejercicios en el modo prueba asociados al cmo c en que el puntaje obtenido es mayor o igual a la mitad del puntaje maximo hasta el mes m)/(total ejercicios en el modo prueba asociados al cmo c hasta el mes m)
Puntaje	vartemp_ptje_p_c	pendiente de la recta que ajusta los datos de la variable [puntaje/puntaje máx] a lo largo del tiempo hasta el mes m
Puntaje	errorpreg_p_c	promedio de errores (variable numerr) sobre todos los ejercicios modo prueba asociados al cmo c hasta el mes m
Puntaje	errorptje_p_c	promedio de la variable [numerr/puntaje máximo] sobre todos los ejercicios modo prueba asociados al cmo c hasta el mes m
Intentos	errorpreg_t_c	promedio de errores (variable numerr) sobre todos los ejercicios modo tarea asociados al cmo c hasta el mes m
Intentos	errorptje_t_c	promedio de la variable [numerr/puntaje máximo] sobre todos los ejercicios modo tarea asociados al cmo c hasta el mes m
Tiempo de Respuesta	tpopreg_t_c	promedio de tiempos de respuesta (variable segundos) sobre todos los ejercicios modo tarea asociados al cmo c hasta el mes m

Tiempo de Respuesta	tpopreg_p_c	promedio de tiempos de respuesta (variable segundos) sobre todos los ejercicios modo prueba asociados al cmo c hasta el mes m
Tiempo de Respuesta	tpoptje_t_c	promedio de la variable [segundos/puntaje máximo] sobre todos los ejercicios modo tarea asociados al cmo c hasta el mes m
Tiempo de Respuesta	tpoptje_p_c	promedio de la variable [segundos/puntaje máximo] sobre todos los ejercicios modo prueba asociados al cmo c hasta el mes m
Tiempo	tpo_ultimo_ejer_c	tiempo transcurrido entre el último ejercicio realizado asociado al cmo c y el mes m
Tiempo	tpoentreejer_c	promedio de los tiempos entre realizaciones de ejercicios asociados al cmo c hasta el mes m
Nº de ejercicios	nejer_t_c	(nº de ejercicios modo tarea asociados al cmo c hasta el mes m) / (nº de meses)
Nº de ejercicios	nejer_p_c	(nº de ejercicios modo prueba asociados al cmo c hasta el mes m) / (nº de meses)
Nº de ejercicios	nejer_c	suma de "nejer_t_c" con "nejer_p_c"

Tabla 2: Variables de desempeño histórico, su descripción y tipo.

Para cada una de estas variables se calcula la métrica KS con su respectiva estimación de desviación estándar. Así, por cada contenido se obtienen 18 valores de la métrica KS, y además se incluye el KS del mes de análisis con el objeto de saber cuánto influye el mes o diferentes períodos del año. Luego, se escogen las variables con mejor KS, lo que es típico de los algoritmos de inducción de árboles de decisión (Friedman, 1977; Araya & Gigon, 1992). Por ejemplo, en el caso de ejercicios de prueba, para el contenido Completar secuencias numéricas y geométricas, el contenido que mejor predice su desempeño es Valor posicional en dígitos medido a través de la variable Puntaje en pruebas (promedio de la variable [puntaje/puntaje máximo] sobre todos los ejercicios modo prueba asociados al cmo "Valor posicional en dígitos" hasta el mes m), con lo que se logra un KS de 37,3% y con una desviación estándar de 3,0 puntos porcentuales. Este KS está medido en la base de validación. En algunas redes se muestra además de la variable con mejor KS, la de segundo mejor KS pero seleccionada dentro de otro tipo de variable (como está especificado en la tabla 2). De esa manera para cada contenido llegan dos ligazones causales, cada una con su KS.

Adicionalmente, en las redes se muestra también justo debajo del KS de la variable principal, el KS cuando esta variable forma parte de un modelo causal del tipo árbol de decisión y el KS de un modelo del tipo support vector machine que combina linealmente variables históricas. En esos dos modelos se utiliza la variable principal que se indica en la ligazón pero acompañada de hasta dos variables adicionales. Estos valores KS indican el potencial de la variable histórica de predecir el desempeño en el contenido donde llega la ligazón al ser complementada esa variable con otras dos adicionales.

Resultados

Los resultados de los algoritmos se presentan como redes. Es una forma visual de representar los hallazgos de manera de ayudar a entender los patrones, tal como lo sugiere (U.S. Department of Education, 2012b), quien destaca la importancia de la actividad de búsqueda de representaciones visuales adecuadas y que denomina "Visual Data Analytics".

La figura 4 muestra la red causal en pruebas, en la que se selecciona para cada contenido aquel contenido histórico con mayor capacidad de predecirlo según la métrica KS. Las líneas entre

contenidos son dirigidas, es decir tienen una punta de flecha que señala la dirección de causalidad. Además, las líneas se dibujan más gruesas a mayor grado de causalidad según el KS. En esta figura se observa la centralidad del eje de Números, cuyos contenidos están de color verde. Esos contenidos predicen el desempeño en los contenidos de Geometría y que están en color celeste, y el de Azar (estadística y probabilidades) que está de color rojo. Esta red tiene las métricas evaluadas en un conjunto de validación, que usa datos estudiante-mes diferentes a los usados para la selección de variables y la construcción de la red. Se observa que el desempeño futuro en los próximos 3 meses en el contenido *resolución de problemas con operaciones* es predicho por el mismo contenido. El signo positivo indica que a mejor desempeño histórico en pruebas en el contenido (medido por *tasa_ptje_p*) mejor es el desempeño futuro. El KS asociado es 35,3 con una desviación estándar de 2,3. Además, si se usa la variable *tasa_ptje_p* para un modelo más completo de árbol de decisión con hasta dos variables adicionales, también se logra un KS de 35,3 lo que está señalado debajo del KS anterior. Y si se usa la variable *tasa_ptje_p* para hacer un modelo support vector machine usando el software WEKA (Bouckaert et Al, 2008) que combina linealmente esa variable con la segunda con mejor KS (no mostrada en la figura) se obtiene un KS de 28.0 con desviación estándar 2,4 que se señala en la tercera fila. Por otra parte el desempeño futuro en los próximos 3 meses en el contenido *cálculo mental y escrito con números naturales* lo predice mejor el desempeño en pruebas en *resolución de problemas con operaciones* con un KS de 37,4 y desviación estándar de 2,2.

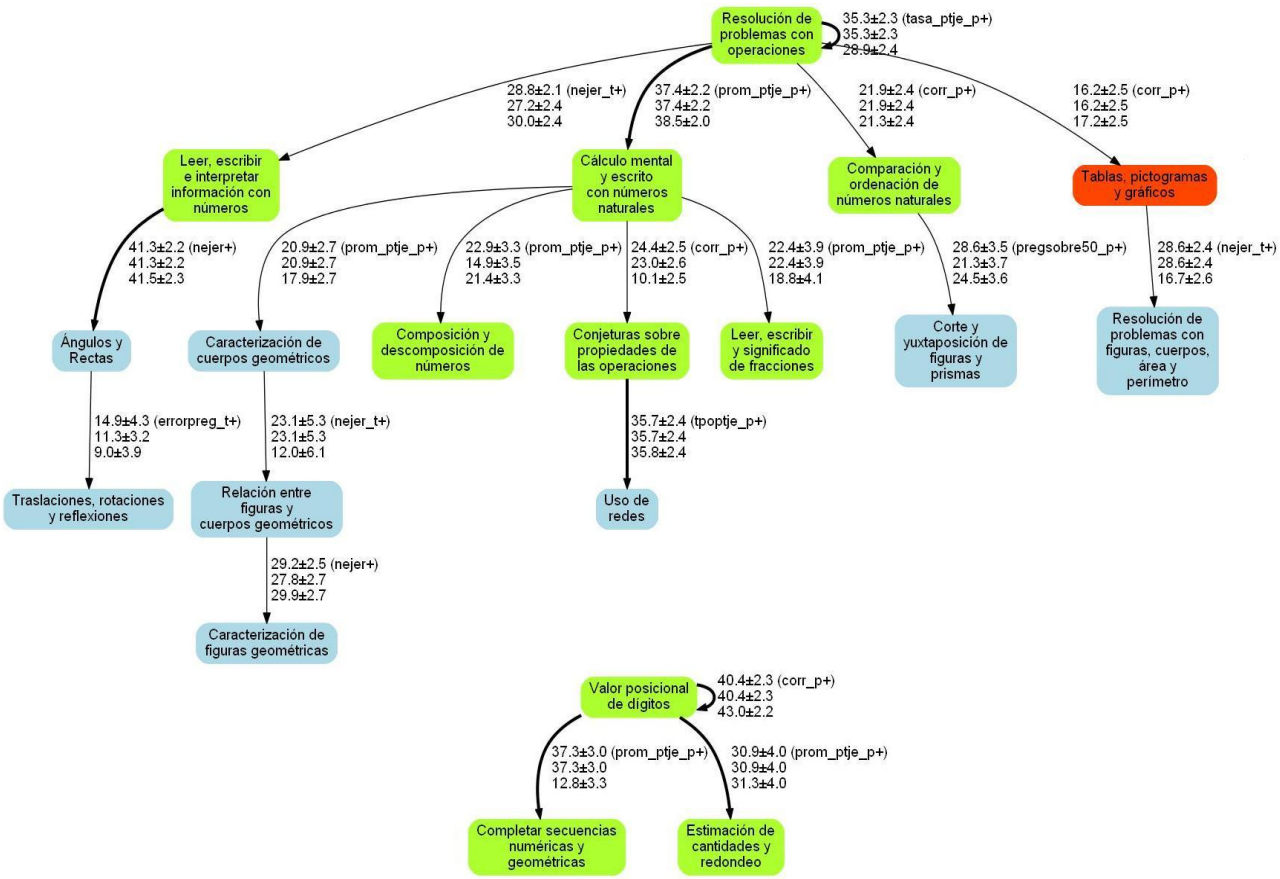


Figura 4: Red de causalidad entre desempeños en contenidos en modo prueba, con KS medidos en base de validación.

Es importante señalar que esta red no tiene que necesariamente coincidir con la red de dependencia lógica o dependencia matemática de los contenidos. Esta es una red de desempeños que depende además de factores psicológicos y sociales, y de cómo los profesores introducen los contenidos del currículo.

La figura 5 muestra una red similar pero construida para los ejercicios tipo tarea. Tal como la red de pruebas, también están indicados los KS en la muestra de validación. Aparece mayor cantidad de contenidos, es decir tiene mayor granularidad, debido al número mucho mayor de ejercicios hechos como tareas. Eso hace que haya más ejercicios por contenidos y más pares alumno-mes no vacíos, y por lo tanto, menos categorías de contenidos sean sacados del análisis. Sin embargo, el grado de predicción (KS) es menor que en las pruebas. En tareas donde hay retroalimentación inmediata y los estudiantes ensayan varias veces, el desempeño histórico no predice tan bien el desempeño futuro. Hay que observar que KS de 0 es cuando no hay evidencia de capacidad de predicción. Por otra parte, hay varias conexiones con KS tales que están a menos de una desviación estándar del cero y otras a menos de dos desviaciones estándares del cero. En esos

casos la dependencia de causalidad es la mejor que existe, sin embargo no es significativa, y por lo tanto no se han dibujado en la red.

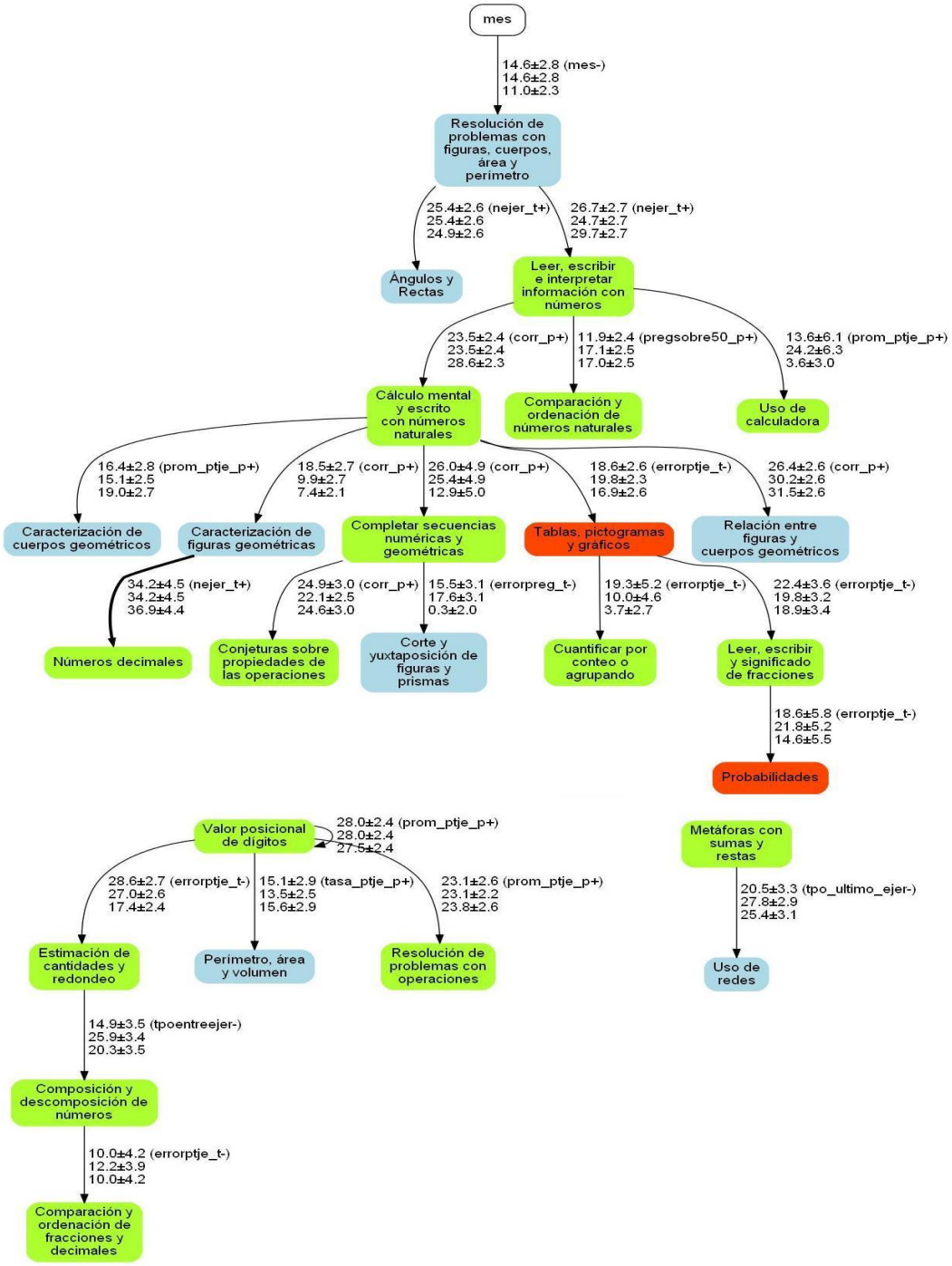


Figura 5: Red de causalidad entre desempeños en contenidos en modo tarea, con KS medidos en base de validación.

Si construimos la red pero con categorías de contenidos exclusivamente dentro de cada eje, dejando fuera del análisis categorías de contenidos de otros ejes, tenemos entonces en la figura 6 las redes causales para el modo prueba.

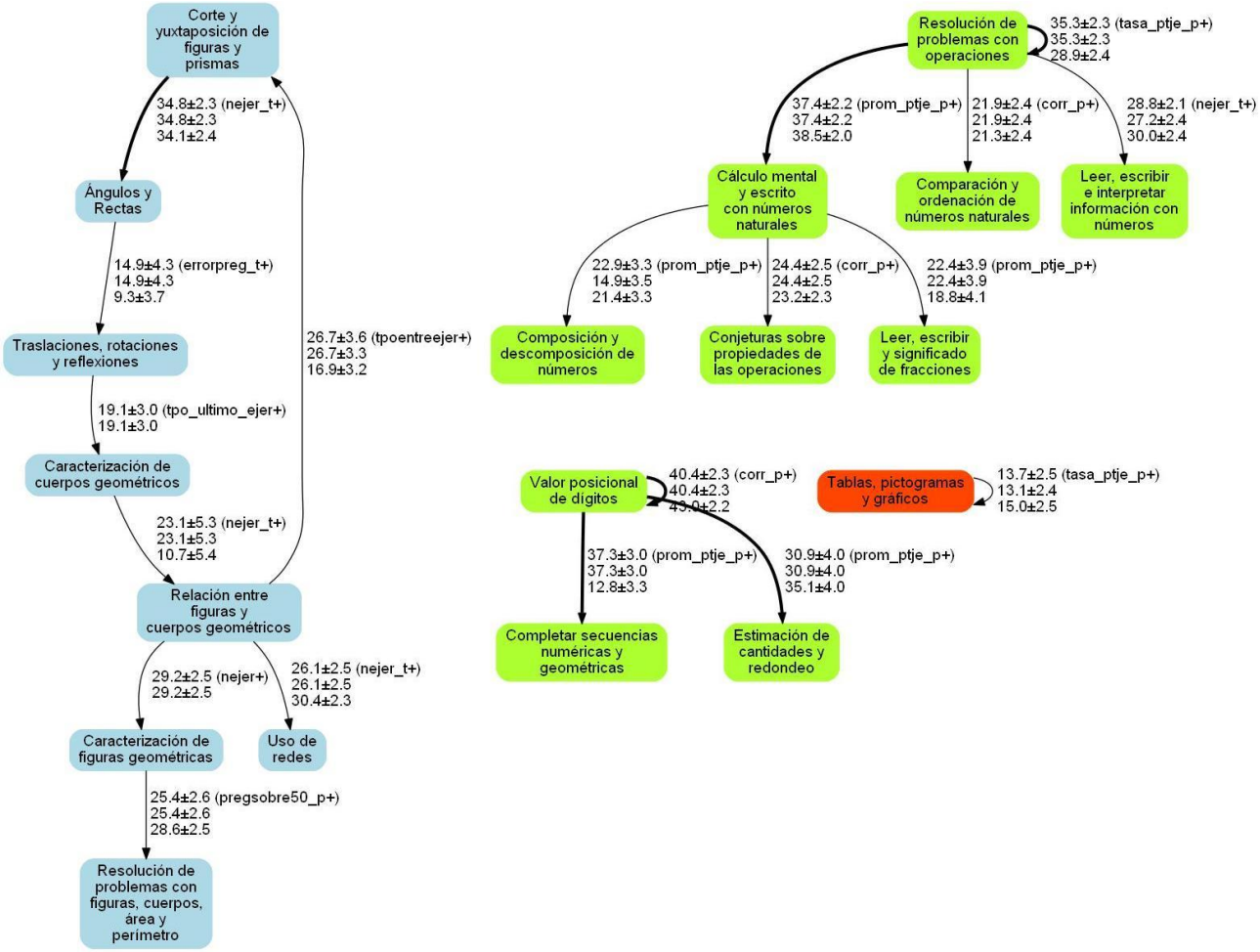


Figura 6: Redes de causalidad entre desempeños en contenidos en modo prueba dentro de cada eje, con KS medidos en base de validación.

Además de la primera variable histórica de desempeño que mejor prediga un contenido es interesante incluir la segunda mejor variable histórica. Así por ejemplo, en la figura 7 se muestra la red de causalidad modo prueba separado por ejes con las dos mejores variables históricas que predicen los contenidos.

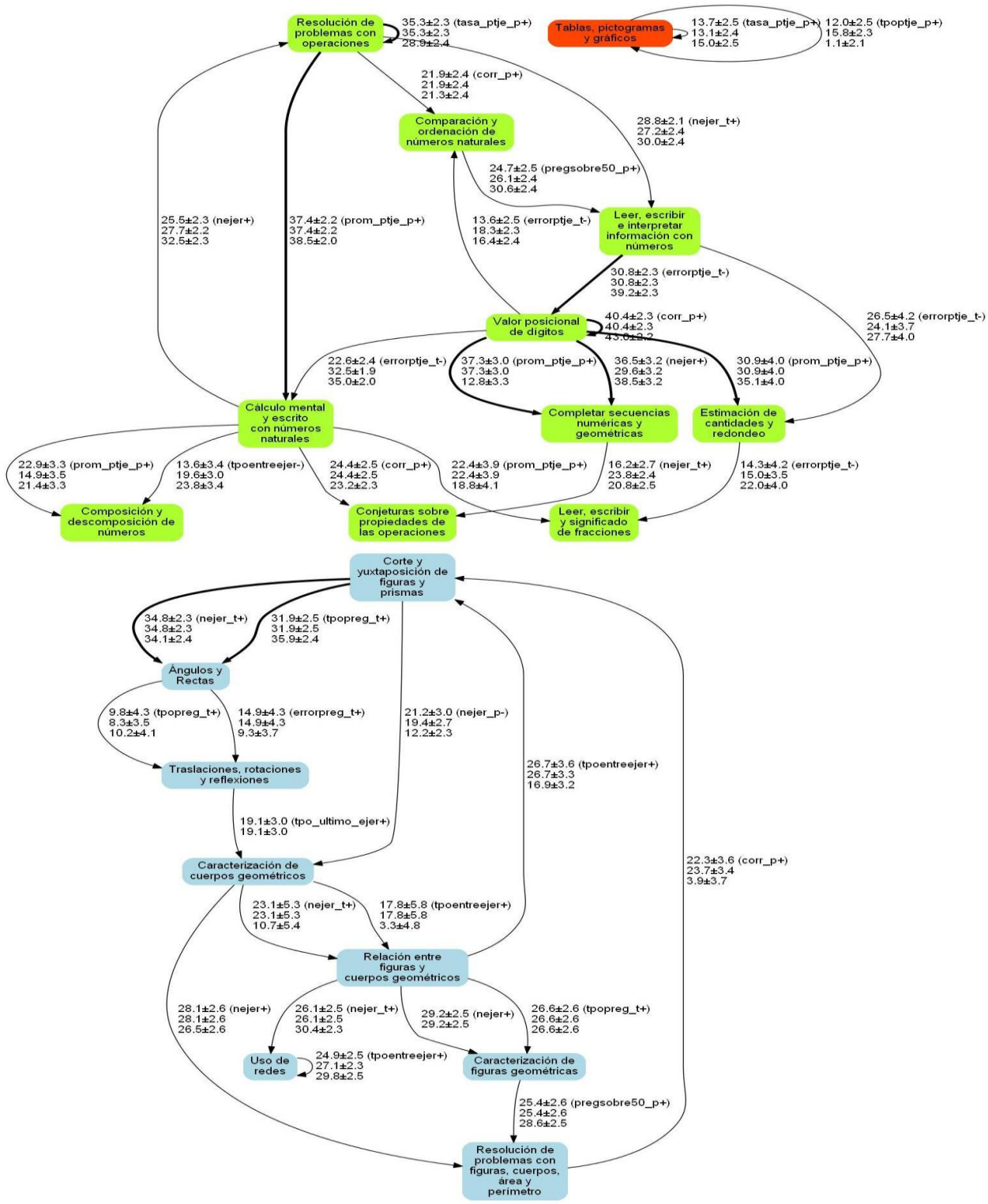


Figura 7: Para cada eje se muestra la red de causalidad entre desempeños en contenidos en modo prueba con los dos desempeños históricos con mejor capacidad de predicción, con KS medidos en base de validación.

La figura 8 muestra las interacciones causales entre los tres ejes: números, geometría y azar. Para cada eje se muestra el mejor impacto del desempeño histórico en los otros dos ejes y de sí mismo, en el desempeño en el futuro cercano del eje. Se observa el fuerte impacto de números sobre sí mismo y los otros dos ejes. Este gran impacto no es solo producto de la dependencia lógica y matemática entre los contenidos y de los procesos psicológicos del aprendizaje, sino que también del tiempo que dedican los profesores a los ejes y las secuencias particulares con que presentan los contenidos.

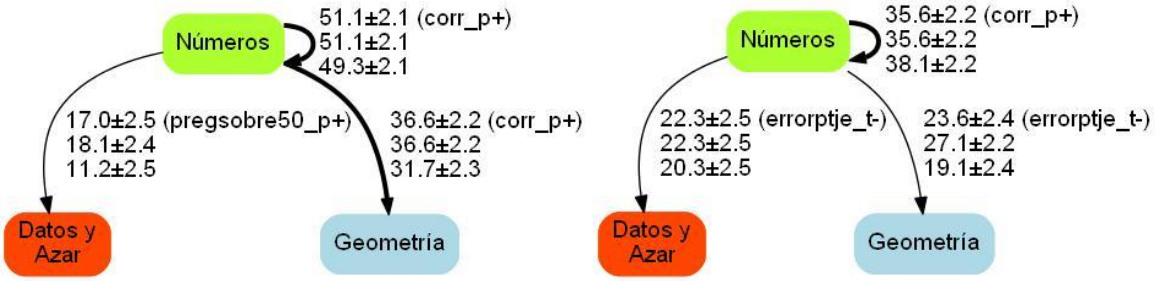


Figura 8: Red de causalidad entre desempeños en los tres ejes del currículo en modo prueba (izquierda) y en modo tarea (derecha), mostrando la influencia causal de cada eje sobre los otros dos y sí mismo, con KS medidos en base de validación.

Adicionalmente a la red, es ilustrativo tener la matriz completa de cómo el desempeño en el futuro cercano en cada contenido puede ser predicho por el desempeño histórico en los contenidos. Así en la figura 9 para cada desempeño futuro descrito en cada fila se muestra en grados de gris la intensidad del KS de cada desempeño histórico descrito en las columnas. Se observa por ejemplo, que el desempeño en el futuro cercano en cálculo menta y escrito con números naturales es predicho mejor por el desempeño histórico en ese mismo contenido y en los desempeños históricos en los contenidos comparación y ordenación de números naturales, completar secuencias numéricas y geométricas, y valor posicional de dígitos.

También es ilustrativo leer la tabla de la figura 9 por columna. Es decir, para el desempeño histórico en un contenido seleccionado ver en qué contenidos puede tener impacto en el futuro cercano. Por ejemplo, de la figura 9 se observa que el desempeño histórico en pruebas en Tablas, pictogramas y gráficos tiene mayor impacto en Ángulos y Rectas que en otros contenidos.

Es muy importante observar que esta tabla no es simétrica. Es decir, el impacto del desempeño pasado en un primer contenido en el desempeño futuro de un segundo contenido, no es el mismo impacto que tiene el desempeño pasado en el segundo contenido en el desempeño futuro del primer contenido. A diferencia de medidas de correlación que miden si los desempeños en ambos contenidos suben o bajan ambos juntos, o al revés si uno sube el otro baja, que sí genera una tabla simétrica, con la medida de causalidad con el KS no se tiene simetría. Esta asimetría refleja el hecho que “si una cosa causa otra, no tiene porqué la segunda cosa causar la primera”.

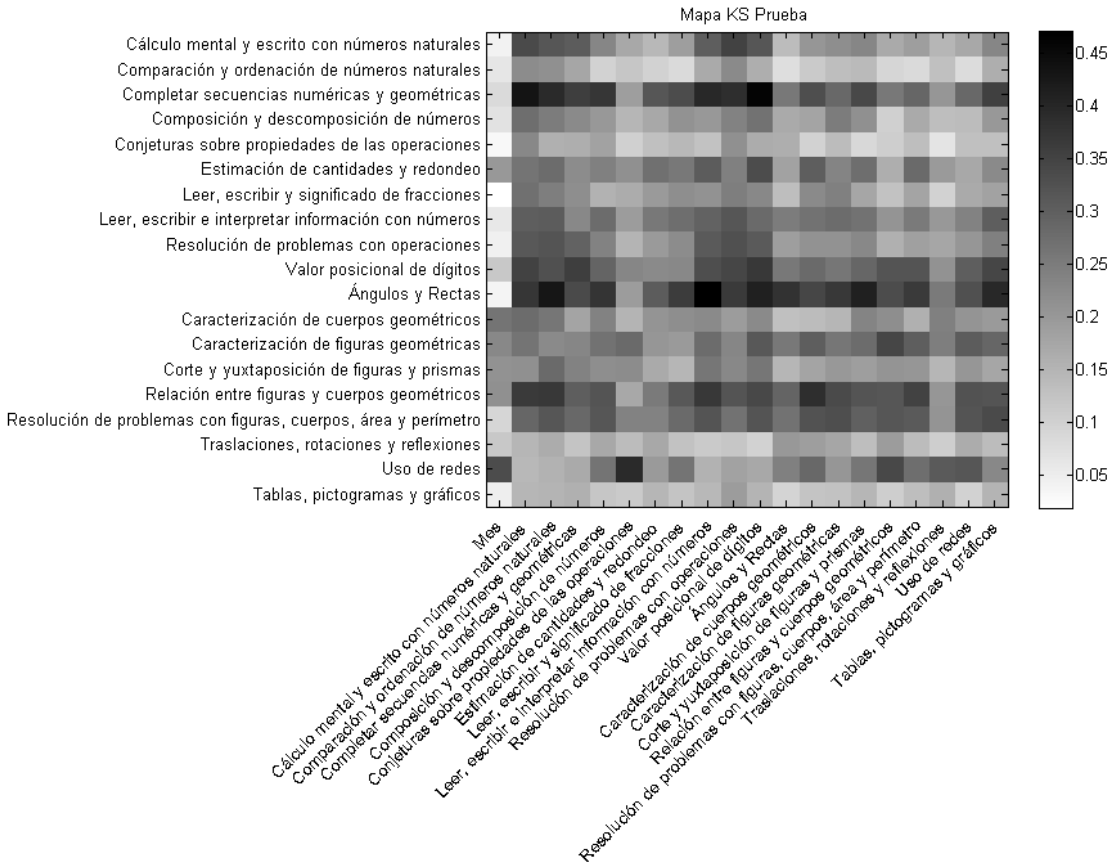


Figura 9: Mapa de KS para pruebas. Para cada contenido su desempeño futuro (filas) se muestran la capacidad de predicción de los contenidos mediante su desempeño histórico (columnas) medidos en la base de validación por el KS en escala de 0 a 1. Se muestran sólo los 17 contenidos con más datos.

Un resumen por ejes puede verse en la figura 10, donde se aprecia la importancia del eje de números.

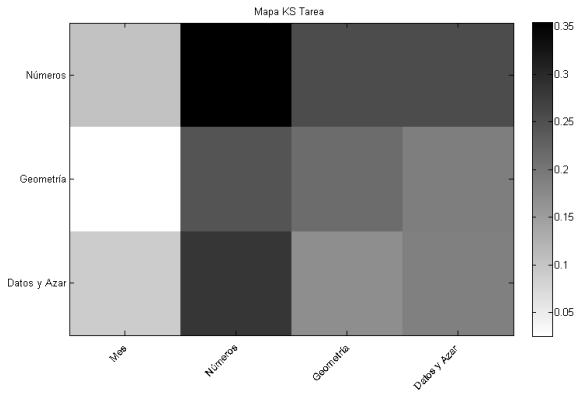


Figura 10: Mapa de KS para tareas por eje. Para cada eje do su desempeño futuro (filas) se muestran la capacidad de predicción de los ejes mediante su desempeño histórico (columnas) medidos en la base de validación por el KS en escala de 0 a 1.

En la figura 11 se muestra la matriz completa de cómo el desempeño en el futuro cercano en cada contenido en ejercicios tipo tarea puede ser predicho por el desempeño histórico en los contenidos. Así en la figura 11 para cada desempeño futuro descrito en cada fila se muestra en grados de gris la intensidad del KS de cada desempeño histórico descrito en las columnas. Así por ejemplo el desempeño en el futuro cercano en Probabilidades es predicho mejor por el desempeño histórico en Leer, escribir y significado de fracción. Por otro lado, leyendo la tabla por columnas podemos observar que el desempeño histórico en Leer, escribir y significado de fracción tiene mayor impacto en el desempeño en el futuro cercano en Probabilidades y también en Números decimales. Es muy interesante observar que el desempeño histórico Números decimales tiene menor impacto en el desempeño futuro de contenidos que el desempeño histórico en Leer, escribir y significado de fracción.

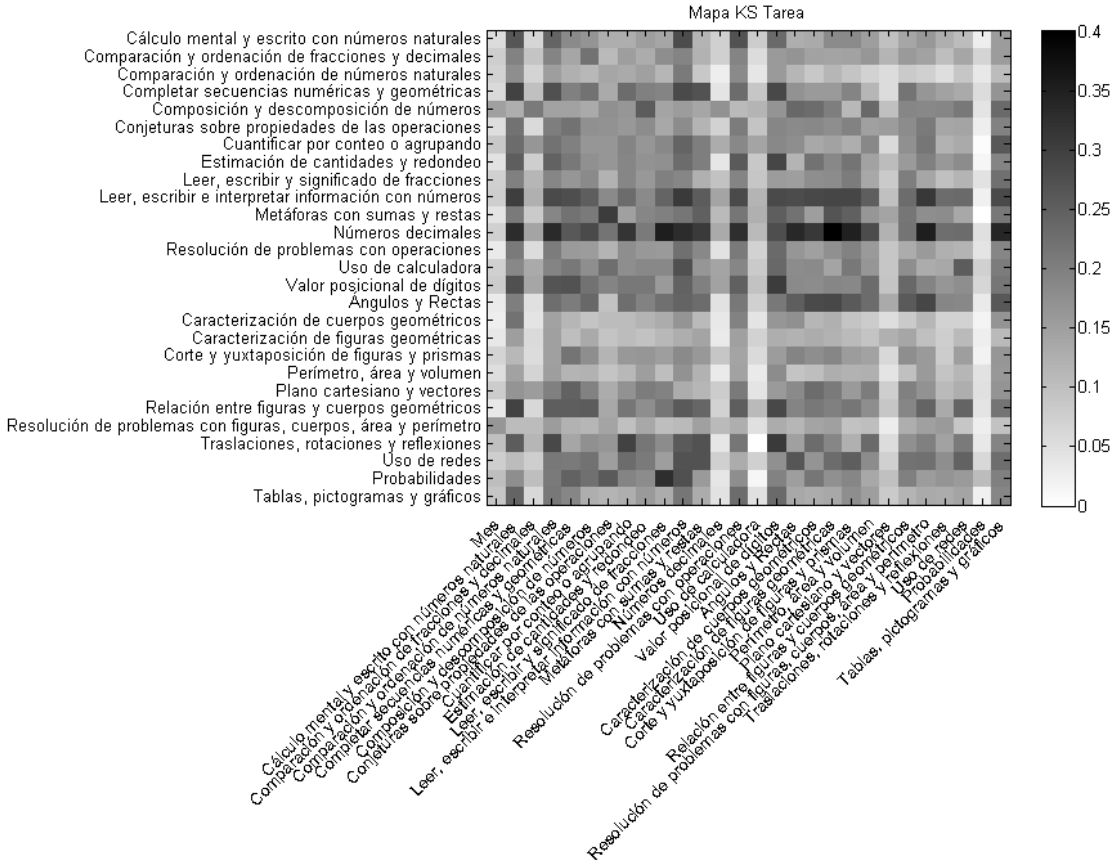


Figura 11: Mapa de KS para tareas. Para cada contenido su desempeño futuro (filas) se muestran la capacidad de predicción de los contenidos mediante su desempeño histórico (columnas) medidos en la base de validación por el KS en escala de 0 a 1. Se muestran sólo los 17 contenidos con más datos.

Conclusiones

El creciente volumen de información, el mayor nivel de detalle de los registros y los nuevos algoritmos de minería de datos están transformando enormemente la investigación científica (Science, 2011) y el desarrollo industrial y comercial (Mckinsey, 2012). Según el reporte McKinsey en el sector educacional todavía la facilidad de captura de datos en compleja, la intensidad de uso de las tic es baja, la cultura de uso de información para la toma de decisiones es débil, y la disponibilidad de información es escasa (se almacena mucho menos información por empleado de lo que lo hacen otros sectores, menos del 10% de lo que lo hace la bolsa de valores y menos de un 20% de lo que lo hace la banca). Sin embargo, con el uso de sistemas en línea y próximamente con el reemplazo de los textos por textos electrónicos, se viene un aumento exponencial de información educacional. Esta explosión de información, capturada con mayor granularidad y frecuencia, con el detalle de las interacciones y aprendizajes segundo a segundo de los estudiantes, está comenzando una etapa de grandes trasformaciones del sector. En este trabajo,

por ejemplo, se analizó la información de un año de ejercitación de una o dos sesiones semanales de 805 estudiantes de cuartos básicos, lo que ya produjo más de medio millón de registros, cada uno con la información de porcentaje de preguntas correctas, número de intentos hasta completar el ejercicio y tiempo de respuestas. En una implementación completa para todas las asignaturas se hubiera producido 10 millones de registros de ejercicios para los cerca de 800 estudiantes. Esta es la cantidad de ejercicios que prontamente registrará una escuela típica de 800 estudiantes. Aún así, esos 10 millones de ejercicios registrados no consideran las tareas hechas en casa ni otro tipo de información como la de redes sociales, imágenes y videos que prontamente se estarán almacenando. Es entonces importante comenzar a aprovechar esta oportunidad y explorar la forma que nos puede ayudar a mejorar los currículos y el aprendizaje. En este trabajo hemos visto que ya con la información que se está almacenando hoy en día en algunas escuelas vulnerables y con metodologías de minería de datos se puede comenzar a medir en forma precisa las interacciones entre contenidos que se producen en el proceso de aprendizaje y encontrar patrones en la causalidad de desempeños en unos contenidos sobre el desempeño en otros.

Una estrategia para analizar la información es el uso de diferentes formas de visualización. Una forma de representación visual adecuada puede ayudar enormemente a detectar patrones y ver interacciones cuando hay una gran cantidad de información. El sistema visual abarca gran parte de la corteza cerebral humana y es una poderosa adaptación biológica para detectar patrones. En la medida que podamos presentar la información en formato visual apropiado lograremos una mayor comprensión de las oportunidades que realmente encierra la información y nos permitirá tomar mejores decisiones, como por ejemplo, hacer cambios curriculares y planificaciones de clases. Por esta razón, es importante explorar algoritmos de visualización de la red de dependencia causal entre contenidos.

Esto significa una enorme oportunidad para el diseño curricular, que cuente con un instrumento detallado para ver las reales interacciones causales de componentes del currículo cuando los estudiantes están aprendiendo, considerando las estrategias y procesos psicológicos que experimentan, y no solo con las estructuras lógicas teóricas entre los contenidos. Así se incorpora un elemento que normalmente ha estado ausente en el diseño curricular y en la planificación de clases: la evidencia empírica del fenómeno sicológico y social del aprendizaje efectivo en las aulas.

El análisis de más de medio millón de ejercicios realizados muestra la importancia y centralidad del eje de números. El desempeño histórico en números predice no solo el desempeño futuro en los contenidos de números, sino que también es un excelente predictor del desempeño en los otros dos ejes. Además, dentro del eje números los contenidos con mayor impacto causal en otros contenidos específicos de números son cálculo mental y escrito con números naturales, resolución de problemas con operaciones y valor posicional de dígitos. En las diferentes redes construidas, tanto en modo prueba como modo tarea, estos tres contenidos aparecen con fuertes conexiones causales a varios otros contenidos. Este hallazgo coincide con el diagnóstico de expertos como los que diseñaron los Common Core Standards en EEUU (Common Core State Standards Initiative, 2011), cuya estrategia es enfocarse en ideas centrales y en particular en los contenidos de números.

Una de las limitaciones de este estudio es el bajo número de cursos analizados. Sin embargo, con los 23 cursos de 13 escuelas urbanas vulnerables puede ya tenerse un primer panorama claro del impacto de las diferentes componentes del currículo en cuarto básico para ese sector. En el futuro compararemos los patrones de causalidad obtenidos con los del año 2012, una vez se complete el año escolar. Además sería importante expandir la medición a más escuelas y cursos de otros sectores, y estudiar si hay diferencias entre tipos de establecimientos. Otro aspecto muy importante a estudiar es analizar las interacciones causales del currículo de matemáticas con el currículo de otras asignaturas

Agradecimientos

Los autores agradecen a Maximiliano Ríos, Director de la Corporación Municipal Lo Prado y a sus directores y profesores de las escuelas de la Corporación; a Jorge Morkoff Representante de la Escuela Básica Santa Rita de Casia de la comuna de La Pintana, a su director y profesores; a Juan Carlos Gutiérrez Gerente General de la Fundación Educacional SECST; a Luis Gianelli Rector de la escuela Complejo Educacional Monseñor Luis Arturo Pérez de la comuna de Pedro Aguirre Cerda y sus profesores; a Raúl Gormaz y Jorge Soto-Andrade por múltiples sugerencias en el análisis de la información; a las profesoras Paulina Jaure, Elena Salinas, Carolina Ramírez, Viviana Palavecinos, Enilse Briones, Cecilia Aravena, Laura Aguayo a cargo de los laboratorios, a Gina Luci que coordinó en el 2011 la actividad en los laboratorios con las jefaturas de las UTP; a Carlos Aguirre, Manuel Bahamondes, Guillermo Campos, Ricardo Jofré por la implementación del sistema en línea, a Patricio Calfucura por soporte técnico; al proyecto Conicyt CIE-05 Centro de Investigación Avanzada en Educación, y al proyecto Fondef TIC EDU TE10I001

Referencias

Araya, R. and Gigon, P. (1992). Segmentation Trees: A New Help for Building Expert Systems and Neural Networks. COMPSTAT92, 1:119-124.

Araya, R. (2000) Inteligencia Matemática. Editorial Universitaria.

Araya, R. & Dartnell, P. (2008) Video Study of Mathematics Teaching in Chile. Procceedings 11th Internacional Conference on Mathematics Education Conference. Monterrey, México.

Araya, R.; Calfucura, P., Jiménez, A.; Aguirre, C.; Palavicino, M.; Lacourly, N.; Soto-Andrade, J. & Dartnell, P. (2010) The Effect of Analogies on Learning to Solve Algebraic Equations. Pedagogies: An International Journal. Special Issue The teaching of Algebra. Volume 5, Issue 3.

Araya, R. & Van der Molen, J. (2012) A mayor ejercitación en línea en matemáticas más aumenta el SIMCE. Por aparecer en Revista Chilena de Educación Matemática RECHIEM.

Banilower, Eric R.; Boyd, Sally E.; Pasley, Joan D. & Weiss, Iris R. (2006) Lessons from a Decade of Mathematics and Science Reform. A Capstone Report for the Local Systemic Change through Teacher Enhancement Initiative.

Bennet, K. & Campbell, C. (2000). Support vector machines: hype or hallelujah? ACMSIGKDD Explorations, December 2000.

Bhatt, R. & Koedel, C. (2011) Large-Scale Evaluations of Curricular Effectiveness: The Case of Elementary Mathematics in Indiana. http://www2.gsu.edu/~ecorrb/index_files/MathCurricula.pdf

Bouckaert, R., Frank, E., Hall, M., Kirkby, R., Reutemann, P., Seewald, A. et al (2008). WEKA manual for version 3-6-0. Hamilton: University of Waikato.

Common Core State Standards Initiative (2011). Common Core State Standards for Mathematics. Bajado de http://www.corestandards.org/the-standards/mathematics

Consortium for Policy Research in Education (2009). School Improvement by Design: Lessons from a Study of Comprehensive School Reform Programs. August 2009.

Friedman, J.H. (1977). A recursive partitioning Decision Rule for Nonparametric Classification. IEEE Transactions on Computers.

Geary, D. (2007). Educating the evolved mind: Conceptual foundations for an evolutionary educational psychology. Charlotte, NC: Information Age Publishing.

Gigerenzer, G. (2000). Adaptive thinking: Rationality in the real world. Oxford: Oxford University Press.

Hadamard, J. (1945). The Mathematician´s Mind. Princeton University Press. New Jersey.

Hattie, J. & Timperley, H. (2007) The Power of Feedback, Review of Educational Research. Vol 77, No. 1, pp. 81-112.

Karpicke & Blunt (2011) Retrieval Practice Produces More Learning than Elaborative Studying with Concept Mapping. Science, 20 January.

McKinsey Global Institute (2011) Big data: The next frontier for innovation, competition, and productivity

Michie, D.; Spiegelhalter, D. & Taylor,C. (1994) Machine Learning, Neural and Statistical Classification. Ellis Horwood.

Mingers, J. (1989). An Empirical Comparison of Selection Measures for Decision Tree Induction. Machine Learning, 3.

National Mathematics Advisory Panel (2008). Report of the Task Group on Instructional Practices.

The National Academies (2001) Knowing What Students Know: The Science and Design of Educational Assessment.

Richland, L.E., Holyoak, K.J., & Stigler, J.W. (2004). Analogy use in eighth-grade mathematics

classroom. Cognition & Instruction, 22(1), 37–60.

Richland, L.E., Zur, O., & Holyoak, K.J. (2007). Cognitive supports for analogies in the mathematics

classroom. Science, 316(5828), 1128–1129.

Roediger, H. & Karpicke, J. (2006) Test-Enhanced Learning: Taking Memory Tests Improves Long Term Retention. Psychological Science.

Science (2011) Special Issue “Dealing with Data” Science, February 2011.

Siegler, R. (1998). Children’s thinking. Upper Saddle River, NJ: Prentice Hall.

Siegler, R. S. & Stern, E. (1998). Conscious and unconscious strategy discoveries: A microgenetic analysis. Journal of Experimental Psychology: General, 127, 377-397.

Siegler, R. & Araya, R. (2005) A computational model of conscious and unconscious strategy discovery. In R. V. Kail (Ed.), Advances in child development and behavior, Vol. 33 (pp. 1-42). Oxford, UK: Elsevier.

The American Statistical Association (2007) Using Statistics Effectively in Mathematics Education Research

The K12 Mathematics Curriculum Center (2005) The Changing Mathematics Curriculum: An Annotated Bibliography

U.S. Department of Education. (2009a) Effectiveness of Reading and Mathematics Software Products Findings From Two Student Cohorts. Institute of Education Sciences National Center for Education Evaluation and Regional Assistance. February 2009

U.S. Department of Education. (2009b) Using Student Achievement Data to Support Instructional Decision Making.

U.S. Department of Education. (2010) Transforming American Education. Learning Powered by Technology.

U.S. Department of Education. (2012a) Improving Mathematical Problem Solving in Grades 4 Through 8.

U.S. Department of Education (2012b) Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics. [draft for public comment] Office of Educational Technology

U.S. Department of Education (2012c) Achievement Effects of Four Early

Elementary School Math Curricula Findings from First Graders in 39 Schools

Van Lehn (1990) Mind Bugs: the Origins of Procedural Misconceptions. Mit Press.