Revista Iberoamericana de Evaluación Educativa

ISSN: 1989-0397

Evaluando el Discurso Docente: Desarrollo de un Modelo de Rasch a Partir de la Evidencia Audiovisual de Profesores Chilenos de Primer Ciclo de Educación Básica en el Área de Lenguaje

Jorge González, David Preiss y Ernesto San Martín

Revista Iberoamericana de Evaluación Educativa 2008 - Volumen 1, Número 2

http://www.rinace.net/riee/numeros/vol1-num2/art9.pdf

Fecha de recepción: 15 de septiembre de 2008
Fecha de aceptación: 6 de octubre de 2008

El propósito de este trabajo es estudiar cuáles son las principales fortalezas y debilidades que muestra el discurso docente de los profesores de primer ciclo –esto es, primero a cuarto básico- que trabajan en colegios públicos en Chile en el área de lenguaje. Para tales efectos, utilizando un Modelo Rasch Poisson, el presente estudio reanaliza los resultados previamente obtenidos por Preiss (En prensa). Específicamente, Preiss analizó los patrones de enseñanza del lenguaje adoptados por profesores de lenguaje de primer ciclo utilizando estrategias de investigación inspiradas en los estudios TIMSS de 1995 y 1999 (Hiebert et al., 2003; Stigler, Gonzales, Kawanaka, Knoll, & Serrano, 1999). Estos estudios se han constituído en un referente para la investigación en sala de clases y en el contexto latino-americano, no sólo Preiss sino que también Loera (2006) ha adoptado parte de las estrategias de investigación de los estudios TIMSS para investigar la naturaleza de la enseñanza. Específicamente, Preiss aplicó estas estrategias para investigar el desempeño de profesores en clases filmadas para el Sistema Nacional de Evaluación Docente, *Docente Mas* (DM, para una descripción del sistema ver Manzi, 2007), utilizado en Chile.

Específicamente, en el ámbito del discurso, el estudio realizado por Preiss mostró que, en todos los niveles de competencia que DM evalúa, las preguntas que hacen los profesores y los seguimientos (follow-up) de éstos a las intervenciones de los estudiantes durante las aperturas de clases estuvieron principalmente focalizados en controlar el flujo de clase o chequear información y, escasamente, en considerar los procesos de pensamiento de los alumnos. Durante los primeros diez minutos de clase, el número promedio de preguntas o seguimientos por profesor que involucraban un componente metacognitivo era menor a tres para todos los niveles de competencia evaluados por DM. Vista la escasa presencia de un componente metacognitivo en el discurso de los profesores, nuestro interés ahora es profundizar el análisis cuantitativo de esta evidencia, utilizando una estrategia de análisis que nos permita estudiar el desempeño comunicativo de los profesores tanto desde el punto de vista de la habilidad de los profesores como de la dificultad que representa para cada profesor cada tipo de pregunta y seguimiento. Así, nuestra hipótesis es que el hecho que las preguntas y follow-up empleados por los profesores estén más o menos basados en ciertas variables de discurso o tareas que el profesor emplea en el aula, a ser definidas más adelante, depende de ambos componentes: la habilidad del profesor y la dificultad de la tarea a ejecutar.

1. Estrategias de análisis

Una forma de cuantificar cuán efectivo es el discurso docente de un profesor es simplemente contar el número de veces que el profesor ejecuta la tarea en cuestión. Por ejemplo, diremos que un profesor alcanza satisfactoriamente la tarea de controlar el flujo de la clase si se le contabilizan un número razonable de preguntas relacionadas a esta tarea (ver más adelante el Cuadro 1 para algunos ejemplos de preguntas). Por el contrario si el profesor ejecuta pocas veces o simplemente no ejecuta la tarea, diremos que no realiza satisfactoriamente la misma. Hay dos aspectos de la naturaleza de estos datos que permiten escoger una metodología estadística de análisis pertinente con los mismos. En primer lugar, se trata de un proceso de conteo: se cuentan las veces que el profesor ejecuta determinadas tareas. Por otro lado, no se le impone al profesor el que realice la mayor cantidad de tareas posibles en el tiempo correspondiente a una clase. Tomando en cuenta estos aspectos, podemos suponer que los datos observados para cada docente provienen de una distribución de Poisson. Ahora bien, parece evidente pensar que si una tarea es fácil, sería contabilizada para la mayoría de los profesores; por otro lado, si un

profesor es muy hábil, se espera contabilizar tareas de distinta dificultad. Es por eso que hacemos depender de dos aspectos el número de veces que el profesor ejecuta una tarea: por un lado, la dificultad de la tarea; por otro lado, la habilidad del docente para realizar dichas tareas.

Un modelo estadístico que captura las características recién mencionadas es el modelo Rasch Poisson de conteos (RPCM, Rasch, 1960b; Van Duijn & Jansen, 1995). Básicamente se asume que, condicionalmente a la habilidad del profesor y la dificultad de la tarea a ejecutar, el puntaje del profesor en una tarea es una variable aleatoria con distribución Poisson y con un parámetro de intensidad que es una función tanto de la habilidad del profesor como de la dificultad de la tarea. La principal ventaja de este modelo es que ambas características pueden ser representados en una escala intervalar común, permitiendo así la comparación entre profesores y, adicionalmente, entre profesores con respecto a la tarea considerada.

2. Metacognición como proxy de la efectividad

Tal como indicamos, uno de los objetivos en este estudio es elucidar cuán difícil le resulta a los profesores incorporar contenidos metacognitivos en su discurso. Un discurso rico en contenido mental es un componente clave de la calidad de la educación. Profesores que practican un discurso metacognitivo producirán estudiantes que serán más capaces de entender las opiniones de otros y las suyas propias (Olson & Astington, 1993). Este tipo de discurso es comúnmente estudiado en el contexto de secuencias IRF (initiate-respond-follow-up) una de las formas dominantes que adopta la comunicación dentro de la sala de clases (Mehan, 1979; Sinclair & Coulthard, 1975). Estas secuencias son regularmente iniciadas por una pregunta del profesor, seguida por una respuesta de un estudiante, y comúnmente cerrada por un follow-up de profesor (Wells, 1999). La evaluación de secuencias IRF es mixta. Algunos autores son críticos de su alta frecuencia en la sala de clases; otros afirman que son, por diseño, claves para alcanzar los fines de enseñanza ya que proveen un cierre -el seguimiento o follow-up- para el contenido de la clase, la cual es colaborativamente construida por el profesor y los estudiantes (Newman, Griffin, & Cole, 1989). Wells (1999) va más allá de estas oposiciones y afirma que no es la ocurrencia de la secuencia IRF la importante sino la manera en la cual es implementada. En efecto, estas secuencias pueden ser usadas para incrementar una respuesta de un alumno o para conectar los contenidos de la clase con su experiencia de vida. Por otra parte, los tópicos pueden ser introducidos indistintamente ya sea por el profesor o los estudiantes, dando forma a distintos tipos de de secuencias IRF. Y, finalmente, los contenidos tratados pueden ser substancialmente diferentes dependiendo de la tarea a mano: introducir contenido o revisar las tareas, por ejemplo (Wells, 1999).

La codificación de secuencias IRF a gran escala es posible, como es ilustrado en los estudios de TIMMS. Por ejemplo, el estudio de videos de TIMMS de 1995 codificó un amplio número de secuencias de enseñanza en tres países diferentes estudiando 90 clases de matemáticas en 8º grado de los Estados Unidos, Alemania y Canadá. La clasificación de expresiones hecha por el estudio coincide estrechamente con las secuencias IRF canónicas, las cuales son consideradas una unidad coherente de intercambio conversacional (Stigler et al., 1999). En los tres países estudiados por TIMMS, las preguntas (las cuales denominaron elicitaciones) con mayor incidencia fueron aquellas en que se pregunta a los alumnos que nombren o declaren un concepto matemático, mientras aquellos que requirieron a los estudiantes describir o explicar fueron menos comunes. Para Estados Unidos el 77% de las elicitaciones iniciales fueron usadas para declarar o nombrar, mientras sólo el 3% fueron usadas para describir o explicar. En el

caso de Japón, 13% de elicitaciones iniciales contenían describir o explicar. Tal como se ve, el uso de metacognición es la excepción, más que la norma en todos estos casos.

3. Preguntas de investigación e hipótesis

Las principales preguntas de investigación del presente estudio son las siguientes:

1. ¿Cuáles son las fortalezas y debilidades del discurso docente de primer ciclo en la enseñanza del lenguaje en las aulas chilenas?
2. ¿Cuán fácil o difícil le resulta a los profesores aquí estudiados realizar preguntas o seguimientos que hagan uso de metacognición?

Nuestra hipótesis es que las fortalezas y debilidades de los docentes dependerán tanto de la habilidad de éstos como del nivel de dificultad de la tarea en ejecución relacionada con el discurso. Así mismo, esperamos que las tareas de carácter metacognitivo serán las de mayor dificultad para los profesores.

4. Muestra de profesores

De un universo de 3740 profesores de enseñanza básica evaluados por DM en 2003, una muestra de 135 profesores fue considerada para los análisis. Los profesores trabajan en establecimientos ubicados en nueve regiones del país, incluyendo el norte, sur y centro. Ochenta y nueve por ciento de los profesores de enseñanza básica en Chile trabajan en zonas urbanas por lo que para la muestra se consideraron sólo estos profesores. De este grupo, se seleccionó a los profesores utilizando muestreo aleatorio simple. El promedio de edad fue 47.09 años (D.E.=8.94). Casos con datos faltantes fueron excluidos del análisis obteniendo una muestra final de 128 profesores.

5. Metodología y materiales

5.1. Estudios de vídeos

Los videos fueron el principal material utilizado en este estudio. Un camarógrafo experto filmó a cada profesor durante 40 minutos. El tiempo de filmación fue acordado con los profesores por medio de los directores de los establecimientos. Se les pidió a los profesores no grabar la primera clase del día ya que comúnmente contienen tareas no relacionadas con instrucciones tales como actividades administrativas, desarrollo de hábitos de higiene, y celebración de cumpleaños u otros feriados. Al menos el 80% de los alumnos del curso debían estar presentes, por lo que se recomendó a los profesores avisar a sus alumnos de antemano sobre la filmación para estimular su asistencia. Los profesores tuvieron que comunicar al camarógrafo de antemano cuales actividades realizarían de manera que el camarógrafo pudiera hacer lo necesario para grabar la clase. Aunque se les pidió a los profesores que realizaran su clase de manera usual, se les pidió que realizaran actividades en las cuales activamente involucrara a sus alumnos e interactuara con ellos. Se les pidió, sin embargo, que incluyeran al menos 10 minutos de exposición de un tema durante la filmación. Actividades pasivas tales como ver una película o leer silenciosamente fueron explícitamente no solicitadas.

5.2. Esquema de codificación del discurso docente

El esquema final de codificación incluyó dos categorías principales para las tareas a evaluar: *preguntas del profesor y "follow-ups" del profesor*. Preguntas son todas las que el profesor hace a toda la clase o individualmente en público. Follow-ups son todas las respuestas públicas hechas por el profesor a las preguntas de los estudiantes. La codificación fue adaptada de Wells (1999) y del estudio de videos de TIMSS 2005 (Stigler, Gallimore, & Hiebert, 2000; Stigler & Hiebert, 1999). Se seleccionaron los códigos que eran más relevantes al contexto Chileno. Las preguntas de los profesores fueron codificadas como Controla, Informa, Implementa y Elabora. Una descripción acabada del proceso de codificación se encuentra en Preiss (2006). Presentamos aquí sólo las definiciones de cada código.

a) *Controla*: El propósito de estas preguntas es regular el flujo de la clase. El profesor evalúa si los estudiantes están siguiendo la clase y organiza algunas tareas.

b) *Informa:* El propósito de estas preguntas es evaluar información. El profesor evalua si los estudiantes saben alguna información o están siguiendo la información que el profesor está dando en clases.

c) *Implementa*: El propósito de estas preguntas es hacer que los estudiantes utilicen habilidades verbales. En particular el profesor entrena al alumno en lenguaje.

d) *Elabora*: El propósito de estas preguntas es pedir a los estudiantes que reflexionen acerca de su trabajo previo y expresen sus creencias. El profesor le da la oportunidad al estudiante de pensar sobre lo que ha hecho o dicho, de colaborar con la clase y de debatir el significado de su trabajo. Usualmente se utilizan el tipo de preguntas ¿por qué? o ¿para qué?

Los "follow up" fueron codificados usando los siguientes códigos

e) *Monosílabo:* El profesor responde una respuesta del alumno con un monosílabo como "uh uh" o "ok" o con una respuesta sin consecuencias para la instrucción.

f) *Repite:* El profesor responde a los aportes de los estudiantes repitiendo parcial o totalmente la respuesta de los estudiantes.

g) *Evalúa:* El profesor realiza una evaluación explícita de la respuesta de los estudiantes. Este código incluye además todos los "follow-up" que tenían una intención de evaluar como por ejemplo "¿estás seguro de tu respuesta?"

h) *Reformula*: El profesor capitaliza sobre un aporte de un estudiante para explorar el contenido.

El Cuadro 1 muestra ejemplos de cada una de las categorías.

CUADRO 1. EJEMPLOS DE PREGUNTAS Y "FOLLOW-UP"

Código	Ejemplo
Control	- ¿Me siguen? - ¿terminaste? ¿alguien más tiene algo que aportar?
Informa	- ¿Qué estoy mostrando?, ¿qué es esto? (indicando una ilustración) - ¿qué es un sustantivo?
Implementa	- ¿qué dice aquí? - ¿qué se debe corregir en la frase?
Elabora	-¿Quién puede explicar porqué esta es la solución correcta al problema? - ¿Cuál crees es la principal enseñanza de esta historia?
Monosílabo	- Uh, uh…*
Repite	- ¡Un gato!*
Evalúa	- ¡Un gato! ¡muy buena respuesta!*
Reformula	¡Un gato! Correcto. Los gatos son mascotas muy comunes. Pero existen otros animales que son mascotas también, por ejemplo los perros. *

*El profesor pregunta por un animal que al mismo tiempo es una mascota. Un estudiante responde "un gato". Entonces el profesor continúa la conversación usando alguno de los "follow-up" que se indican.

6. ANÁLISIS DE DATOS

Sea Y_{ij} la variable aleatoria que denota el número de veces que el profesor i realiza la tarea j $(i=1,...,n\ ;\ j=1,...,k)$. Por ejemplo, $y_{i5}=10$ significa que el *i*-ésimo profesor utilizó 10 veces un monosílabo como follow-up.

El modelo Rasch Poisson (Rasch, 1960b; Van Duijn & Jansen, 1995) asume que la probabilidad que el profesor i tenga y_{ij} ejecuciones de la tarea j se modela como una distribución de Poisson condicionalmente a la habilidad θ_i del profesor y a la dificultad β_j de la tarea de forma que $(Y_{ij}\,|\,\theta_i,\beta_j)\ \square\ Poisson(\exp(\theta_i-\beta_j))$. Las diferencias entre individuos son representadas por medio de las habilidades individuales. Cada profesor es caracterizado por una y sola una habilidad, lo cual no implica que dos o más profesores obtengan la misma estimación de habilidad. Así, si el profesor *i* resulta ser más hábil que el profesor *i´* se tendrá que $\hat{\theta}_i>\hat{\theta}_{i'}$. Similarmente, si la tarea *j* es más difícil de ejecutar que la tarea *j´* se tendrá que $\hat{\beta}_j>\hat{\beta}_{j'}$.Como se mencionó anteriormente, la principal ventaja de este modelo es la capacidad de localizar en una misma escala intervalar habilidades y dificultades, haciendo posible no sólo una comparación directa entre profesores, sino también entre profesores con respecto a la tarea considerada. La especificación del modelo se completa asumiendo que la distribución de habilidades es normal con media cero y varianza σ^2.

Es pertinente indicar que técnicamente se trabaja con un modelo estadístico que se obtiene después de integrar el efecto aleatorio (habilidades individuales). A partir de este modelo, se obtienen los estimadores de las dificultades y de la varianza σ^2. Usando la técnica de empirical Bayes, se generan las habilidades individuales, las cuales se utilizan como estimadores de las habilidades de los docentes. Ahora bien, la representación simultánea habilidad-dificultad *sólo* es válida cuando las habilidades se consideran como parámetros, al igual que las dificultades. Puesto que los estimadores de los parámetros de este último modelo son muy cercanas a las que se obtienen con los estimadores calculados en este trabajo, es posible utilizar la comparabilidad. Para detalles, ver Lindsay, Clogg y Greco (1991) y del Pino, San Martín, González & De Boeck (2008).

Un total de ocho variables, representando ocho tareas distintas fueron utilizadas en el estudio. Las variables pueden ser clasificadas en dos subcategorías como se muestra en el Cuadro 2.

CUADRO 2. VARIABLES UTILIZADAS EN EL ESTUDIO

Nombre de la variable
Tipos de Pregunta
1 Control
2 Información
3 Implementación
4 Elaboración
Tipos de "Follow-up"
5 Monosílabo
6 Repite
7 Evalúa
8 Reformula

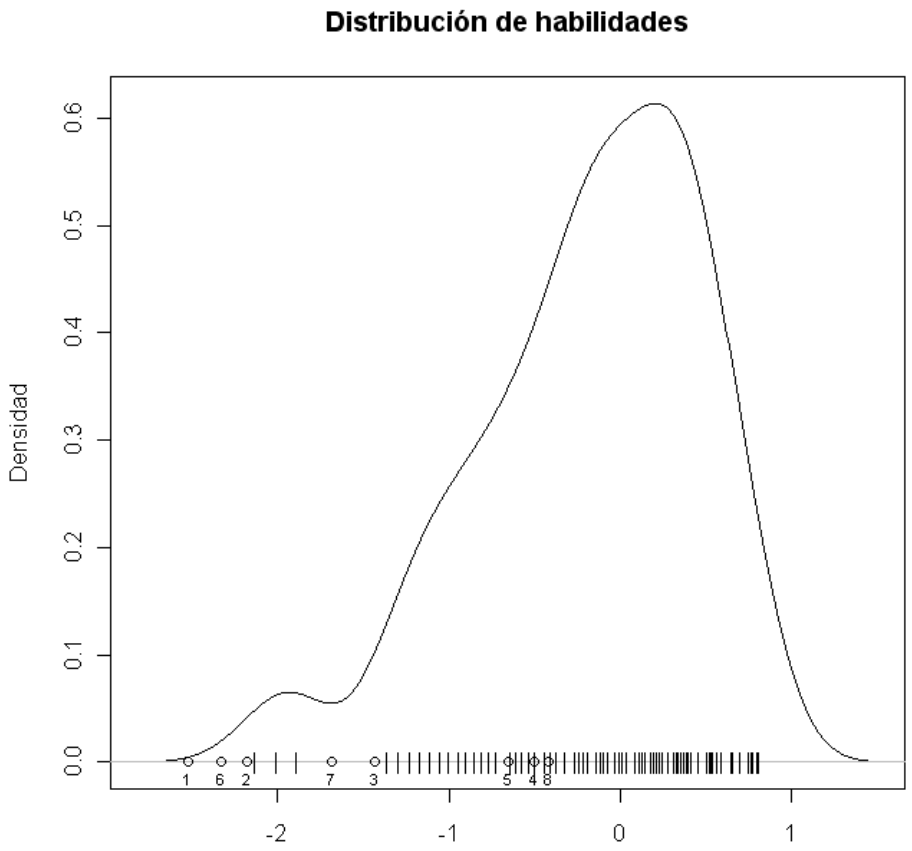
7. RESULTADOS

El Cuadro 3 muestra las dificultades estimadas para los parámetros relacionados con las tareas que definen el discurso docente. En primer lugar se observa que todas las estimaciones son estadísticamente significativas (*p<0000.1*). Se observa que la tarea más difícil para los profesores es *reformula* ($\hat{\beta}_8$=-0.42) mientras que la más sencilla es *control* ($\hat{\beta}_1$=-2.52). La varianza estimada es $\hat{\sigma}^2$=0.47 (*SE=0.04, p<0.0001*). Dado que existen 83 estimaciones distintas de habilidad para los 128 casos considerados (recuerde que algunos profesores pueden recibir la misma estimación de habilidad), es más ilustrativo graficar la distribución de habilidades en lugar de listar dichos 83 valores. La Figura 1 muestra la distribución de habilidades de los profesores evaluados, cuyos valores puntuales están representados por barras y, adicionalmente, en una misma escala, la estimación de dificultades de las tareas representadas por círculos. La estimación de habilidades muestra que una gran porción de profesores realiza satisfactoriamente todas las tareas. El Cuadro 4 muestra el porcentaje de profesores que realiza satisfactoriamente cada una de las tareas que definen el discurso docente.

CUADRO 3. ESTIMACIÓN Y ERROR ESTÁNDAR DE LAS DIFICULTADES DE LAS TAREAS

Parámetro	Estimación	SE
β_1	-2.5179	0.04061
β_2	-2.1757	0.04372
β_3	-1.4353	0.05389
β_4	-0.5008	0.07640
β_5	-0.6570	0.07169
β_6	-2.3250	0.04226
β_7	-1.6857	0.04984
β_8	-0.4203	0.07901

FIGURA 1. DISTRIBUCIÓN DE HABILIDADES Y DIFICULTADES ESTIMADAS. BARRAS REPRESENTAN HABILIDADES Y CÍRCULOS DIFICULTADES ESTIMADAS.



CUADRO 4. PORCENTAJE DE PROFESORES CUYA HABILIDAD SOBREPASA LA DIFICULTAD DE LA TAREA CONSIDERADA

Tarea	Porcentaje
Tipos de Pregunta	
1 Control	100%
2 Información	100%
3 Implementación	96.09%
4 Elaboración	72.66%
Tipos de "Follow-up"	
5 Monosílabo	78.13%
6 Repite	100%
7 Evalúa	96.09%
8 Reformula	68.75%

Se observa por ejemplo que todos los profesores están por sobre la tarea más sencilla, es decir, todos los docentes evaluados controlan el flujo de la clase de manera satisfactoria. De la misma forma, el 100% de los docentes hace preguntas para evaluar información en los estudiantes y repite parcial o totalmente las respuesta de los estudiantes a modo de follow-up. Se concluye entonces que las principales fortalezas de los profesores evaluados se sitúan en estas tres variables de discurso. De la figura también se aprecia que 5 profesores logran realizar sólo 3 de las tareas de manera satisfactoria (en la figura se aprecian 3 barras dado que 3 de los 5 profesores obtuvieron la misma estimación de habilidad). Estos 5 profesores están por sobre el promedio de edad, siendo uno de ellos el de máxima edad.

Sólo un 68.75% de los profesores logra realizar satisfactoriamente la tarea más difícil asi como un 72.66% tiene éxito con la tarea que sigue en grado de dificultad. Las principales debilidades del discurso docente entonces se sitúan en las tareas de carácter más metacognitivo, aquí codificadas como Reformula y Elabora.

8. DISCUSIÓN

En relación a las preguntas de investigación que motivaron este estudio, podemos concluir que las principales fortalezas de los docentes evaluados se relacionan con: el control y la preocupación que muestra el docente para evaluar si los estudiantes están siguiendo la clase, la capacidad de evaluar la información que manejan los alumnos o si estos están siguiendo la información que el profesor está dando en clases, y el hecho de responder a las intervenciones de los alumnos repitiendo parcial o totalmente la respuesta de los estudiantes. Dentro de las debilidades podemos concluir que los docentes no ejecutan de manera satisfactoria aquellas tareas con un marcado carácter metacognitivo, tales como capitalizar sobre una intervención de un alumno para explorar el contenido (Reformular) o evaluar explícitamente las respuestas de los estudiantes (Evaluar). Se desprende de estos resultados que resulta difícil para los profesores aquí estudiados realizar preguntas o seguimientos que hagan uso de la metacognición, un hecho lamentable tomando en cuenta la incidencia que tiene en los aprendizajes y desarrollo de habilidades de los estudiantes el contar con profesores que practican un discurso metacognitivo. Estos hallazgos confirman lo encontrado en el estudio previo de Preiss (En prensa) en que las preguntas y follow-up que hacen los profesores estuvieron enfocadas principalmente en el control del flujo de la clase y chequeo de información. Así mismo, el hecho que las tareas consideradas de mayor carácter metacognitivo -elaboración y reformula- hayan sido el tipo de pregunta y follow-up más difícil de alcanzar para los profesores, respectivamente, coincide con el bajo número promedio de preguntas o follow-up por profesor incorporando un componente metacognitivo encontrado por Preiss.

En términos metodológicos, es importante señalar los siguientes aspectos: en primer lugar, los estimadores tanto de las habilidades de cada profesor, como de las dificultades de las tareas son, respectivamente, funciones monótonas del número de veces que un profesor realiza una determinada tarea, como del número de veces que una determinada tarea aparece en la muestra de profesores consideradas en este estudio. Por lo tanto, usando las proporciones mencionadas, es posible, por un lado, comparar a los profesores entre ellos y, por otro lado, las tareas entre ellas. Cabe entonces preguntarse cuál es el valor metodológico de utilizar un modelo estadístico más complejo como el modelo Rasch Poisson de conteo.

La respuesta a esta pregunta es que, al utilizar dicho modelo, es posible *poner en una misma escala tanto las habilidades de los profesores, como las dificultades de las tareas.* Esta representación simultánea ha sido explotada en la sección 5 y es la que permite interpretar las habilidades individuales en términos de las subtareas. Es importante mencionar acá que los modelos tipo Rasch (uno de los cuales es el utilizado en este estudio) han sido desarrollados para modelar estadísticamente las diferencias individuales. En efecto, Rasch (1960a) ha hecho eco de algunas críticas venidas del campo de la psicología (Skinner, 1956) en el sentido de que esta última sólo se preocupaba de analizar y reportar resultados a nivel poblacional, siendo que el interés sustantivo de los análisis psicológicos eran a nivel individual. Rasch (1960a) desarrolló un par de modelos que explicaban las observaciones a partir de dos tipos de parámetros: aquéllos que representan a las tareas o ítemes (llamados parámetros de dificultad) y los que representan a los individuos (llamados parámetros de habilidad) (para más detalles, ver Mouchart & San Martin, 2003).

Adicionalmente es importante mencionar que esta representación simultánea habilidades-dificultades asume que la distribución de probabilidad que genera las dificultades es la verdadera (en este estudio, una distribución normal con varianza desconocida); dicho en otros términos, que no hay un problema de mala especificación de dicha distribución. El ordenamiento entre profesores y subtareas podría cambiar si dicha distribución no es del todo correcta; pero evidentemente esto requiere estimar la distribución de probabilidad que genera dichas habilidades –estimarla, pues se trata de la distribución de objetos que no son directamente observables. Estadísticamente, se trata de un modelo mucho más complicado, y que será estudiado por los autores en una posterior publicación.

REFERENCIAS BIBLIOGRÁFICAS

Del Pino, G., San Martín, E., Gonzalez, J., y De Boeck, P. (2008). On the relationships between sum score based estimation and joint maximum likelihood estimation. *Psychometrika, 73*(1), pp. 145-151.

Hiebert, J., Gallimore, R., Garnier, H., Givvin, K. B., Hollingsworth, H., Jacobs, J., *et al.* (2003). *Teaching Mathematics in Seven Countries: Results From the TIMSS 1999 Video Study,* (No. NCES 2003–013 Revised). Washington, DC: U.S. Government Printing Office, o. Document Number)

Lindsay, B., Clogg, C. C., y Grego, J. (1991). Semiparametric Estimation in the Rasch Model and Related Exponential Response Models, Including a Simple Latent Class Model for Item Analysis. *Journal of the American Statistical Association, 86*(413), pp. 96-107.

Loera, A. (2006). *La práctica pedagógica videograbada. [Teaching practices videotaped].* México, DF: UPN.

Manzi, J. (2007). Evaluación Docente: Antecedentes, resultados y proyecciones [Electronic Version].

Seminario Evaluación Docente en Chile: Fundamentos, experiencias y resultados. Retrieved January 2, 2008, from http://mideuc.cl/seminario/presentaciones.html

Mehan, H. (1979). *Learning lessons: social organization in the classroom*. Cambridge, MA: Harvard University Press.

Mouchart, M., y San Martin, E. (2003). Specification and identification issues in models involving a latent hierarchical structure. *Journal of Statistical Planning and Inference, 111*(1-2), pp. 143-163.

Newman, D., Griffin, P.,y Cole, M. (1989). *The construction zone: working for cognitive change in school*. New York: Cambridge University Press.

Olson, D. R., & Astington, J. W. (1993). Thinking about thinking: Learning how to take statements and hold beliefs. *Educational Psychologist, 28*(1), pp. 7-23.

Preiss, D. D. (En prensa). The Chilean instructional pattern for the teaching of language: a video-survey study based on a national program for the assessment of teaching. *Learning and Individual Differences*.

Rasch, G. (1960a). On General Laws and the Meaning of Measurement in Psychology. *Proceedings of the Berkeley Symposium on Mathematical Statistics and Theory of Probability* (4), pp. 321-333.

Rasch, G. (1960b). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.

Sinclair, J. M., y Coulthard, M. (1975). *Towards an analysis of discourse : the English used by teachers and pupils*. London: Oxford University Press.

Skinner, B. F. (1956). A Case History in Scientific Method *The American Psychologist, 11*, pp. 221-233.

Stigler, J. W., Gallimore, R., y Hiebert, J. (2000). Using video surveys to compare classrooms and teaching across cultures: Examples and lessons from the TIMSS video studies. *Educational Psychologist., 35*, 87-100.

Stigler, J. W., Gonzales, P., Kawanaka, T., Knoll, S., y Serrano, A. (1999). *The TIMSS Videotape Classroom Study: methods and findings from an exploratory research project on eighth-grade mathematics instruction in Germany, Japan, and the United States* (Research and Development Report No. NCES 1999-074). Washington, D.C.: U.S. Government Printing Officeo. Document Number)

Stigler, J. W., y Hiebert, J. (1999). *The Teaching Gap. Best ideas from the world's teachers for improving education in the classroom*. New York: The Free Press.

Van Duijn, M. A. J., y Jansen, M. G. H. (1995). Modeling Repeated Count Data - Some Extensions of the Rasch Poisson Counts Model. *Journal of Educational and Behavioral Statistics, 20*(3), pp. 241-258.

Wells, G. (1999). *Dialogic inquiry: towards a sociocultural practice and theory of education*. New York: Cambridge University Press.