

# V

## ¿Es la evaluación de profesor por los estudiantes un incentivo perverso?

*Is student evaluation of teacher a wrong incentive?*

Pontificia Universidad Javeriana, Colombia

### **Yuri Gorbaneff**



Periodista Internacional del Instituto Estatal de Relaciones Internacionales de Moscú del Ministerio de Relaciones Exteriores de la URSS - MGIMO (1978). Maestría en Economía de la Pontificia Universidad Javeriana, Bogotá (2000). Profesor del Departamento de Administración, Facultad de Ciencias Económicas y Administrativas, Pontificia Universidad Javeriana. Miembro fundador del grupo de investigación Grupo de estudios de la dirección estratégica y organizaciones - GEDEO, categoría B en Colciencias. Autor de libros y artículos sobre la enseñanza de ciencias administrativas, incentivos, costos de transacción y sistemas de clasificación de pacientes en salud.

## Resumen

**Objetivo:** Diseñar un modelo formal de la estructura de incentivos creada por la evaluación del profesor por los estudiantes y hacer su prueba empírica.

**Método:** Un juego dinámico de información perfecta en tres etapas. Observaciones realizadas en un curso introductorio de administración en 2013 - 2014. Se analiza la correlación entre la nota promedio y la evaluación del profesor por los estudiantes.

**Resultados:** El modelo muestra equilibrios múltiples y existencia de la correlación en casi todas las celdas.

**Conclusiones:** La evaluación del profesor por los estudiantes se correlaciona con las calificaciones de los alumnos lo que conduce a la inflación de notas.

**Palabras clave:** Educación en administración, incentivos, evaluación del profesor por los alumnos.

## Abstract

**Objective:** To design a game theoretic model of incentive structure created by the student evaluation of teacher and to make an empirical test of the model. **Method:** I use game theory to build a dynamic perfect information game in three stages. To prove the model, I collect observations in the Management introductory course in 2013 – 2014 and make test of correlation between the average grades per class and the student evaluation of teacher. **Results:** The model shows multiple equilibria and existence of the correlation in almost every cell. In some cells the correlation is desirable, in other cells, undesirable. **Conclusions:** The student evaluation of teacher is correlated with the grades and results in grade inflation.

**Keywords:** Management education, incentives, teacher evaluation.

## Introducción

Evaluación de profesor por estudiantes (EPE) es un instrumento de uso común para medir el desempeño de los docentes en las universidades. Los aspectos de la enseñanza, que se evalúan en una importante universidad privada de Bogotá (Colombia), incluyen estrategias pedagógicas que usa el profesor, su dominio del contenido del curso, la interacción profesor - estudiantes, el método de evaluación de estudiantes utilizado por el profesor.

Se le pide al estudiante llenar el formulario de evaluación de profesor individualmente en línea. Las evaluaciones individuales se agregan para cada clase de cada curso. La evaluación final muestra los números promedio de la clase y el número de estudiantes en la clase que llenó el formulario, pero no sus nombres. La evaluación final se entrega a la gerencia del departamento académico y a los profesores implicados. EPE se usa para decidir el ascenso de profesores de tiempo completo. EPE es un insumo fundamental para decidir la continuidad de los profesores a tiempo parcial.

Los gerentes de universidades asumen que los estudiantes son personas maduras autónomas, se esfuerzan por aprender, son capaces de diferenciar al buen maestro de uno malo. Además, son veraces e imparciales, capaces de hacer una buena evaluación para un buen profesor y mala evaluación para un mal profesor, independientemente de las notas altas o bajas obtenidos en el curso.

El supuesto se basa en la teoría constructivista. Según los constructivistas, el estudiante no es un recipiente vacío para ser llenado con el conocimiento. El alumno es un participante activo en el proceso de educación. Aprende los conceptos no memorizando su definición sino cuando se enfrenta a los problemas y trata de resolverlos. Según Vygotsky (Vygotsky, 1986):

El tema central del desarrollo durante la edad escolar es la transición de recuerdo primitivo y atención involuntaria a los procesos mentales superiores de la atención voluntaria y la memoria lógica. Atención, previamente involuntaria, se vuelve cada vez más dependiente del pensamiento propio del niño; memoria mecánica cambia a la memoria lógica guiada por el sentido, y ahora puede ser utilizada deliberadamente por el niño. Se puede decir que la atención y la memoria se convierten en "lógica" y voluntaria, puesto que el control de una función es un homólogo de la conciencia de esta función. Intelectualización de una función y el control voluntario de la misma son dos momentos del mismo proceso de la formación de las funciones mentales superiores (pp. 166 - 167).

Si el alumno es un participante activo, tiene el derecho de evaluar el desempeño de profesores. La teoría de la agencia introduce un poco de escepticismo en este cuadro ideal. Según la teoría, hay dos tipos de actores en un intercambio económico: principal y agente. El principal contrata al agente para realizar cierto trabajo. Con el fin de agilizar el proceso de trabajo, el principal empodera al agente. Esta relación presenta aspectos problemáticos.

En un primer lugar, es la información oculta, o selección adversa: el agente sabe más acerca de la labor realizada, que el principal. En un segundo lugar, el problema de la acción oculta, o el riesgo moral. Como la acción del agente no es directamente observable por el director, el agente puede desinformar al principal. En un tercer lugar, el resultado de la acción del agente también depende de medio ambiente, y no es fácil aislar su esfuerzo del impacto de factores externos. En un cuarto momento, el principal y el agente están maximizando sus propias funciones de utilidad que tienen poco en común (Gorbaneff, Torres, & Cardona, 2009). Con el fin de alinear los intereses de los actores, el principal diseña incentivos.

Un ejemplo de este tipo de incentivos es EPE. El gerente, actuando como un principal, diseña incentivos para hacer al profesor a entregar lo mejor de sí a los estudiantes. Pero esta relación es más compleja e implica otro nivel: la relación docente - alumno, donde el profesor juega el rol de principal, y el estudiante, el de agente. El estudiante, de acuerdo con la teoría, trata de maximizar su función de utilidad. ¿Cuáles son los componentes de su función de utilidad? Las buenas calificaciones y el aprendizaje (Ewing, 2012), menos el esfuerzo necesario para conseguirlos:

$$U = g + r - e$$

Donde:

U es la satisfacción de estudiante

g las notas

r el aprendizaje

e el esfuerzo

El incentivo incorporado en EPE, orienta al profesor a conseguir la satisfacción de estudiantes porque un estudiante satisfecho hará una buena EPE. ¿Qué puede hacer el profesor para elevar la satisfacción de estudiantes? En primer lugar, puede crear un atractivo ambiente de aprendizaje y ser riguroso

en la evaluación. Alternativamente, puede facilitar el proceso de aprendizaje para ahorrar el esfuerzo de estudiante, y ser generoso en las notas. La primera estrategia no es una tarea fácil. La forma más sencilla de tener estudiantes satisfechos es por medio de las altas notas. Esta es la razón porque EPE es objeto de debate. Algunos investigadores (Johnson, 2003) encuentran correlación entre las calificaciones y EPE y llegan a la conclusión de que profesores intercambian altas calificaciones por altas puntuaciones EPE. Ewing (Ewing, 2012), analiza la correlación entre EPE y las calificaciones esperadas. Él utiliza notas esperadas en vez de las definitivas porque supone que estudiantes no saben su calificación definitiva en el momento de hacer EPE. Él encuentra un efecto positivo de las notas esperadas sobre EPE. Langbein (Langbein, 2008), Griffin (Griffin, Hilton, Plummer, & Barret, 2014), Braga (Braga, Paccagnella, & Pellizzari, 2014), Schneider (Schneider, 2013) reportan el mismo efecto. Según los autores, profesores que ponen altas calificaciones a sus estudiantes, obtienen puntuaciones más altas en EPE. Albu (Albu & Badea, 2012) no reportan resultados empíricos sobre el efecto de las calificaciones sobre EPE, sino la opinión de los profesores que consideran que sí existe tal relación. Vásquez (Vásquez & Gabalán, 2006) indaga sobre los atributos de profesor que tienen la mayor importancia para el alumno a la hora de hacer EPE, y encuentra que dos primeros son: si el profesor facilita el aprendizaje y si su sistema de calificación es justa (p. 237). Sus hallazgos confirman los de Ewing (Ewing, 2012).

Greenwald y Gillmore (Greenwald & Gillmore, 1997) explican la correlación notas – EPE con la ayuda de la teoría de atribución, según la cual, alumnos interiorizan la responsabilidad de su éxito pero externalizar la responsabilidad de su fracaso. Cuando estudiantes obtienen altas calificaciones, piensan que ellos las merecen. Cuando estudiantes obtienen bajas calificaciones, piensan que el responsable es el profesor. Por eso el profesor tiene el incentivo incorrecto de “comprar” altos puntajes en EPE con altas calificaciones. Vásquez (Vásquez & Gabalán, 2006) recomienda contrastar EPE con la nota obtenida por el alumno porque malos estudiantes tienden a acusar al profesor de su bajo rendimiento (p. 238).

Otros investigadores (Marsh & Roche, 2000) piensan que EPE es un buen mecanismo. Reconocen la existencia de pequeña correlación calificaciones - EPE, pero no aceptan que esta correlación se debe a la generosidad en notas. Marsh propone una explicación alternativa que se reduce a lo siguiente. Un buen profesor motiva a los estudiantes a trabajar duro. El trabajo duro se traduce en el aprendizaje, y el aprendizaje, en la satisfacción intrínseca y en

las altas calificaciones. Un estudiante satisfecho pone altos puntajes en EPE. Por eso la correlación notas – EPE no es preocupante no tiene nada que ver con la indulgencia e inflación en las notas.

El mismo hecho ( correlación entre las notas y EPE) tiene dos interpretaciones opuestas. ¿Por qué ocurre esta situación? En primer lugar, porque la comunidad académica carece de un modelo conceptual que explique la estructura de incentivos creada por EPE para el profesor y el estudiante.

En segundo lugar, porque la evidencia empírica acerca de la correlación notas – EPE es mixta. La falta de la evidencia sobre la correlación tiene que ver con la complejidad de la enseñanza - aprendizaje. También tiene que ver con el método utilizado por los investigadores. Artículos publicados analizan una serie de cursos que dictan varios departamentos académicos. Esta variedad es una fuente de ruido porque el investigador hace comparaciones entre varios cursos y departamentos académicos que tienen diferente atractivo, métodos pedagógicos, políticas de calificaciones y diferente reputación entre los estudiantes.

Las cosas se ponen peor cuando investigadores incluyen en sus muestras pequeños cursos electivos compuestos por una o algunas clases. En pequeños cursos electivos estudiantes tienen una fuerte motivación y saben de antemano el nombre del profesor. Estos elementos hacen que el supuesto de aleatoriedad de la matrícula no se sostiene y la correlación notas – EPE se hace espuria (Marsh & Roche, 2000). Hay algunas excepciones como Griffin (Griffin, Hilton, Plummer, & Barret, 2014), que controlan por el departamento académico, docente y curso.

El objetivo de este trabajo es diseñar un modelo de teoría de juegos que explica la estructura de incentivos creada por EPE, y hacer una prueba empírica del modelo.

El presente trabajo se diferencia de la literatura anterior que explora la correlación notas - EPE. Se pretende crear un modelo que puede ser el eslabón perdido para explicar la correlación. Los datos fueron recogidos en una universidad y en un sólo curso compuesto por varias clases. El muestra se limita a un solo curso para minimizar los efectos de la variedad de métodos pedagógicos utilizados en diversos departamentos académicos y cursos. Los datos fueron recogidos en una universidad privada colombiana. Presente trabajo es el primer estudio de este tipo en América Latina, hasta donde el autor sepa.

El trabajo hace una contribución práctica porque ayuda a diseñar incentivos en las universidades. En un sentido más amplio, el documento también contribuye a la literatura sobre el pago por desempeño.

## **Método**

La situación se modela como un juego dinámico de información perfecta en tres etapas. En la primera etapa la Naturaleza decide si el profesor es bueno o malo. Un buen profesor es una persona inteligente profesional y emocionalmente. Un mal profesor es un buen profesional que carece de la inteligencia emocional.

En la segunda etapa el profesor selecciona la estrategia de indulgencia o rigor en las notas. Cuando el profesor selecciona la estrategia de indulgencia, persigue la satisfacción de estudiantes. Cuando el profesor selecciona la estrategia de rigor, persigue el aprendizaje de estudiantes. Estas estrategias se formularon de acuerdo con Marsh (Marsh & Roche, 2000).

En la tercera etapa el alumno decide qué papel va a jugar: de un pragmático o de un alumno motivado. Cuando el estudiante selecciona la estrategia pragmática, hace altas calificaciones su prioridad. Cuando el estudiante selecciona la estrategia de motivación, hace el aprendizaje su prioridad.

Los pagos representan las utilidades del profesor y estudiante. La utilidad del profesor es EPE, la utilidad del estudiante es cierta mezcla de la nota y el aprendizaje. El juego se resuelve mediante la inducción hacia atrás con el fin de encontrar la combinación óptima de las estrategias del profesor y el estudiante.

Después de analizar las estrategias de equilibrio, se hace una prueba empírica del modelo para buscar una posible correlación notas – EPE. Esta parte del trabajo es un estudio de campo (McGrath, 1994) que utiliza datos secundarios y puede ser clasificado como un estudio documental. Este tipo de estudio permite realizar observaciones directas de un sistema sin ningún tipo de perturbación. Uno de los reverses de este método es poca posibilidad de generalizar porque la muestra puede no ser representativa de la población. En el presente caso, la muestra es intencional porque sólo se observaron los alumnos de curso introductorio de administración. Sin embargo, la inscripción de los alumnos en las clases es aleatoria, y la asignación de los profesores en cada clase también lo es. Los estudiantes no conocen la asignación de profesores cuando seleccionan sus cursos y clases.

En 2014 se contactó el Departamento de administración de la Escuela de Negocios de una universidad privada en Bogotá. Se obtuvieron y analizaron los datos recogidos en tres semestres en 2013 - 2014 en el curso de introducción a la administración. El curso está abierto a todos los estudiantes de la Universidad, siendo obligatorio para las carreras de administración y contabilidad.

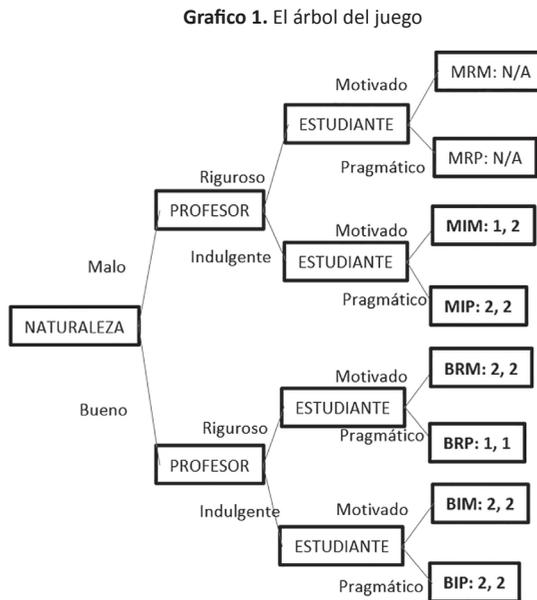
El curso tiene de 11 a 13 clases paralelas de 35 a 40 estudiantes cada una. Es impartido por 8 - 10 profesores de tiempo completo y de tiempo parcial que utilizan el mismo programa, textos, casos y exámenes.

Siguiendo Marsh (Marsh & Roche, 2000), como unidad de análisis fue usada una clase. Fue obtenido el promedio de la nota para cada clase del curso y EPE para cada clase. El número total de observaciones en la base de datos es 38. El análisis consistió en la correlación bivariada que se realizó en SPSS.

Siguiendo Langbein (Langbein, 2008), se utilizaron notas definitivas debido a que EPE tiene lugar después de los exámenes finales, la misma semana cuando se publican las calificaciones definitivas.

## Resultados

El Gráfico 1 reproduce el árbol de juego. El juego tiene múltiples equilibrios: al equilibrio de Nash se puede llegar sólo en un juego repetitivo utilizando las estrategias mixtas. La correlación notas – EPE existe casi siempre, independientemente de la calidad del profesor. Abajo se analiza cada celda del juego:



MIM: en esta celda se observa una falta de correlación, lo cual es deseable. Los estudiantes motivados pueden tener incentivos para reportar la baja calidad de la enseñanza cuando su profesor es malo.

MIP: en esta celda se observa la correlación indeseable. Un mal profesor es indulgente. Un mal estudiante recibe una alta nota y recompensa al profesor con un alto puntaje de EPE.

BRM: en esta celda se observa la correlación deseable. El profesor es bueno y riguroso. El estudiante motivado aprende, obtiene altas calificaciones y pone alta EPE.

BRP: en esta celda se observa la correlación indeseable. El profesor es bueno y riguroso. Pone una baja nota al estudiante pragmático. El estudiante, en represalia, pone un bajo puntaje de EPE.

BIM: una correlación deseable. Buen maestro enseña bien y es indulgente. Él pone altas calificaciones a los estudiantes motivados. Estudiante motivado aprende y pone alta EPE.

BIP: una correlación indeseable. El profesor es bueno e indulgente. Él pone alta calificación al estudiante pragmático. Estudiante pragmática es agradecido y , además, reconoce la calidad de profesor. Él pone altas puntuaciones en EPE.

Para hacer una prueba empírica del modelo, se buscó la correlación notas – EPE. La Tabla 1 muestra una correlación moderada entre EPE y el promedio de nota por clase. La correlación es significativa a 5%.

**Tabla 1**  
*Correlación entre la nota promedio y EPE para cada clase*

Variables e indicadores /Variables		Nota promedio	EPE
Nota promedio	Correlación de Pearson	1	,409*
	Sig. (bilateral)		,011
	Numero de observaciones	38	38
EPE	Correlación de Pearson	,409*	1
	Sig. (bilateral)	,011	
	Numero de observaciones	38	38

\* La correlación esta significativa a 0,05 (2 colas)

## Discusión

El mecanismo de EPE no permite distinguir entre profesores buenos y malos porque las puntuaciones de EPE son difíciles de interpretar. La correlación notas – EPE puede significar varias cosas. Puede ser señal de una alta calidad de profesores, un resultado deseable sugerido por Marsh: un buen profesor enseña bien y el estudiante obtiene altas calificaciones. La correlación puede indicar la indulgencia en las calificaciones pero no puede revelar si el profesor indulgente es bueno o malo.

La falta de correlación puede indicar dos cosas diferentes. Puede ser señal de una baja calidad del profesor malo pero indulgente, que proporciona un estudiante motivado. La falta de correlación también puede ser señal de alta calidad de profesor que proporciona un mal estudiante. Estos resultados están de acuerdo con los de García (García, 2014) quien muestra que el impacto de EPE sobre la calidad de la docencia es nulo (p. 13).

En todas estas situaciones la existencia de correlación conduce a la inflación de notas. Este resultado es similar al obtenido por Schneider (Schneider, 2013). Como ilustración, se puede comparar las estrategias de profesor BR (bueno y riguroso) y BI (bueno e indulgente). La estrategia BR es riesgosa para el profesor porque él profesor va a poner una baja nota al estudiante pragmático y correrá el riesgo de obtener una baja puntuación en EPE en represalia. Por el contrario, la estrategia BL no presenta tales riesgos: tanto en la celda BIM como en BIP el profesor obtiene alta puntuación en EPE. Para el profesor, ser riguroso es una estrategia riesgosa, mientras que ser indulgente es una estrategia dominante. La misma escuela de negocios donde se recogieron los datos ilustra la inflación de nota. Antes de 2011 la escuela aplicaba el 3,2 como el promedio de nota necesario para permanecer en la escuela. En 2011 la nota promedio se elevó a 3,4 sin ningún efecto sobre la calidad de los graduados, medida por los exámenes oficiales estatales Saber Pro.

La falta de equilibrio en estrategias puras sugiere que el maestro y el alumno no siempre juegan la misma estrategia. Ellos saben que juegan un juego repetido y utilizan estrategias mixtas: a veces el maestro es riguroso, a veces es indulgente; A veces el estudiante aparece motivado; a veces se porta como un pragmático. Este resultado es similar al obtenido por Griffin (Griffin, Hilton, Plummer, & Barret, 2014) quien encontró que los profesores cambian su comportamiento cuando enseñan en un rango de clases.

El gerente universitario ayuda a dar forma a la mezcla de estrategias utilizadas por el profesor. Si el administrador presta mucha atención a EPE,

presiona a los profesores hacia la estrategia de indulgencia. Si el administrador no presta mucha atención a la EPE, señala que el profesor es libre de ser exigente con los estudiantes.

¿Qué se puede hacer para resolver el problema de incentivos creados por EPE? El juego muestra que solo los estudiantes motivados tienen incentivos correctos para reportar la calidad de profesores, mientras que los estudiantes pragmáticos no tienen tal incentivo. Este resultado sugiere que una manera de superar el problema consiste en eliminar el anonimato del estudiante en EPE. El profesor y el gerente deben tener acceso no solo a datos agregados, sino a las EPE individuales formuladas por cada estudiante, y relacionar estos resultados con las notas del estudiante. En la práctica eso va a significar tener en cuenta únicamente las opiniones de los estudiantes motivados que aprueban el curso. Para evitar los problemas éticos, estudiantes pueden utilizar seudónimos y la gerencia puede hacer EPE después de que las calificaciones finales se entregan a los estudiantes.

Mi trabajo tiene algunas limitaciones. El simple juego de una sola jugada no representa la naturaleza repetida de la interacción profesor - alumno y no permite modelar la probabilidad de utilizar cierta estrategia. Un futuro trabajo en este campo puede estar basado en un juego repetido con estrategias mixtas. Puede ser útil un juego de señales que permite detectar una verdadera naturaleza de un participante a pesar de sus declaraciones.

La parte empírica de mi investigación se centra exclusivamente en la correlación notas – EPE. Un futuro investigador puede reconocer otros factores que afectan EPE. Por ejemplo, puede ser útil analizar respuestas a preguntas sobre algunos aspectos del curso, relacionar datos cualitativos con puntuaciones cuantitativas, pedir a los estudiantes explicar, cómo y por qué ellos calificaron a su profesor de la manera que lo hicieron. Estas formas de investigación pueden proporcionar más información en profundidad, pero no serán posibles antes de eliminar el anonimato estudiantil en EPE.

**Yuri Gorbaneff**

Pontificia Universidad Javeriana, Colombia  
email: yurigor@javeriana.edu.co

Recibido: 07 de julio de 2015  
Aceptado: 22 de diciembre de 2015

## Referencias

- Albu, G., & Badea, M. (2012). The Student – Evaluator of the Teacher. A Case Study in the University of Ploiesti, Romania. *Procedia - Social and Behavioral Sciences*, 69(24), 2257-2263.
- Braga, M., Paccagnella, M., & Pellizzari, M. (2014). Evaluating students' evaluations of professors. *Economics of Education Review*, 41(August ), 71-88.
- Ewing, A. (2012). Estimating the impact of relative expected grade on student evaluations of teachers. *Economics of Education Review*, 31(1), 141-154.
- Gorbaneff, Y., Torres, S., & Cardona, J. (2009). El concepto de incentivo en administración. Una revisión de la literatura. *Revista de Economía Institucional*, 11(21), 73 - 91.
- Greenwald, A., & Gillmore, G. (1997). No pain, no gain? The importance of measuring course workload in student ratings of instruction. *Journal of Educational Psychology* 89, 743-751.
- Griffin, T., Hilton, J., Plummer, K., & Barret, D. (2014). Correlation between grade point averages and student evaluation of teaching scores: taking a closer look. *Assessment & Evaluation in Higher Education*, 39(3), 339-348.
- Johnson, V. (2003). *Grade inflation: a way out*. Springer.
- Langbein, L. (2008). Management by results: Student evaluation of faculty teaching and the mis-measurement of performance. *Economics of Education Review*, 27(4), 417-428.
- Marsh, H., & Roche, L. (2000). Effects of grading leniency and low workload on students' evaluations of teaching: Popular myth, bias, validity, or innocent bystanders? *Journal of Educational Psychology*, 92(1), 202-228.
- McGrath, J. (1994). *Methodology matters: doing research in the behavioral and social sciences*. Urbana: University of Illinois.
- Muñoz-Cantero, J., Ríos-de-Deus, M., & Abalde, E. (2002). Evaluación Docente vs. Evaluación de la Calidad. *Revista Electrónica de Investigación y Evaluación Educativa*, 8(2), 103 - 134.
- Schneider, G. (2013). Student Evaluations, Grade Inflation and Pluralistic Teaching: Moving from Customer Satisfaction to Student Learning and Critical Thinking. *Forum for Social Economics*, 42(1), 122-135.
- Vásquez, F., & Gabalán, J. (2006). Percepciones de estudiantiles y su influencia en la evaluación del profesorado. *Revista Electrónica de Investigación y Evaluación Educativa*, 219 - 245.
- Vygotsky, L. (1986). *Thought and language*. Cambridge, Mass.: The MIT Press.