



V Jornadas de Sociología de la UNLP
y
I Encuentro Latinoamericano de Metodología de las Ciencias Sociales
“Cambios y continuidades sociales y políticas en Argentina y la región
en las últimas décadas. Desafíos para el conocimiento social”
La Plata, 10, 11 y 12 de diciembre de 2008

SILVIA ISABEL RADULOVICH
UNIVERSIDAD NACIONAL DE SALTA
sradulov@unsa.edu.ar

PONENCIA: EVALUACIONES NACIONALES: TAN LEJOS Y TAN CERCA.

RESUMEN

Trabajo referido a una investigación sobre las funciones políticas de las evaluaciones nacionales en dos países latinoamericanos: Brasil y Uruguay estableciendo comparaciones respecto de estas funciones en Argentina, Bolivia y Chile, ya investigadas.

Las políticas de reforma educativa de los '90 en Latinoamérica toman a la educación como vehículo de efectivización de un modelo social y a las evaluaciones nacionales como vías de justificación de las mismas a través de sus diferentes funciones políticas: primarias, secundarias y terciarias. Estas evaluaciones presentan algunas diferencias en el plano morfológico y algunas semejanzas en el normativo mientras que la discursividad sobre la evaluación -coherente con sus concretas aplicaciones- se articula con el discurso tecnocrático produciendo reducciones que nos muestran la batalla socio-cultural y política por los sentidos que nos interpela en la acción educativa y social.

EVALUACIONES NACIONALES: TAN LEJOS Y TAN CERCA...

... o cómo crear ficciones educativas...

Silvia Isabel Radulovich¹

Universidad Nacional de Salta

sradulov@unsa.edu.ar

Introducción

Cuando escuchamos o leemos acerca de evaluaciones de sistemas educativos² lo primero que se nos aparece son imágenes de cifras, porcentajes y distintas representaciones gráficas de cuadros estadísticos. Estas así llamadas 'evaluaciones' afectan de diverso modo la cotidianeidad institucional en ámbitos educativos, y es inherente a nuestra responsabilidad social como investigadores del campo de lo social, abordar específicamente un fenómeno que nos enfrenta con la problemática de imposición de una *lógica* ajena que afecta la propia naturaleza formativa de la educación institucionalizada trasladando tan cerca nuestro espacios simbólicos y criterios de la lógica económica globalizada. Imposición que llega inclusive a condicionar la práctica áulica en distintos países. Fenómenos éstos ya estudiados ampliamente³, y que muestran la progresiva trastocación del ciudadano con derecho a la educación, en *cliente que consume educación* al mejor postor. Las políticas de reformas educativas promueven evaluaciones macro en los países de la Región realizando una contribución notable a este trastocamiento, más allá de la retórica que las justifica. Nos interpela para pensar y reflexionar sobre los *sentidos* involucrados.

Interesa específicamente abordar esta temática pues yendo un poco más profundo, presenta un punto (no el único, claro) que articula la *formulación macro* y la *realización micro* de políticas educativas, sobre todo en relación con los efectos -o consecuencias- individuales e institucionales de las evaluaciones nacionales (altas o bajas implicancias⁴).

¹ Docente regular de Evaluación Educativa, Facultad de Humanidades, UNSa. Directora Proyecto N° 1694 del Consejo de Investigación de la Universidad Nacional de Salta: *Evaluaciones nacionales: comparación de sus funciones políticas en Brasil y Uruguay. Relaciones con Argentina, Bolivia y Chile*.

² A lo largo de este Trabajo, se utilizará el significante 'evaluación' porque es de uso habitual en distintos documentos y discursos, pero consideramos que no se trata de evaluaciones sino estrictamente de mediciones que darán lugar a valoraciones, *a posteriori* de algún análisis.

³ House (1998: 27-38, 59-64); Pini (2003:99-102, 115-138), entre otras referencias investigativas. Véase el enfoque de rendición de cuentas de McMeekin (2006:19-43) en Corvalán y McMeekin.

⁴ Prueba de altas implicancias: da lugar a alguna forma de castigo para aquellos que obtienen puntajes bajos o alguna forma de recompensa para aquellos con puntajes altos. (Bracey, 2002:23) . Las de bajas implicancias no afectan a sujetos ni a instituciones.

¿Qué hay debajo de la alfombra?

A.

Las políticas públicas de Reforma de la educación desde los '90 en Latinoamérica presentan entre otras una marcada tendencia a unir (a.) la implantación de una propuesta curricular nacional centralizada con (b.) la organización de un sistema nacional de evaluación. Esta articulación proviene de los acuerdos de los gobiernos con los agencias internacionales de crédito (F.M.I., B.M., BID) y se van plasmando políticamente, junto a otros acuerdos, a través de la Organización de Estados Iberoamericanos. Otra vía política es el denominado Proceso de Cumbres de las Américas, tan significativa y pormenorizadamente analizada por Feldfeber y Saforcada (2005). Las autoras muestran cómo *...en las acciones de cooperación multilateral.....se observa en todo el continente incluida América del Norte... el énfasis puesto en la promoción de las nuevas tecnologías de la información y la comunicación (TICs) y en el desarrollo de sistemas estandarizados de evaluación de los resultados... [y cómo] se desarrollan políticas centradas en la utilización de tests estandarizados por lo general vinculados a las iniciativas desarrolladas por la OCDE.* (pág.27, 28). Citan a Kuhlen quien afirma que *...no se sorprendan cuando se enteren que el Banco Mundial impone como condiciones de otorgamiento de préstamos para la educación, la exigencia de que todos los alumnos sean sometidos a un test estandarizado, oportunamente vendido por empresas cuya sede central se encuentra en los Estados Unidos (2004:21)*⁵.

Cuestión que Laurence Wolff, consultor del BID desde hace 11 años -y antes empleado del Banco Mundial durante 22 años- explicita con claridad al afirmar que los resultados de las evaluaciones a nivel nacional se utilizan para *...controlar y supervisar el sistema educativo en todos los niveles...*, incluso especifica que su aplicación es *...una condición necesaria para establecer metas cuantitativas, evaluando las ventajas comparativas de las estrategias de asignación de recursos* (Wolff, 1998: 4,21). Se trata de un experto en costos que investigando en 2007 ha encontrado que *...los costos de evaluación tal y como se realizan en la Región (por ejemplo pruebas muestrales, o evaluaciones censales a determinados grados) no representan una carga pesada para los presupuestos educativos...*

⁵ Citado de Kuhlen, L. (2004) "La política educativa del gobierno de Bush. Miles de niños olvidados" en *Boletín de la Secretaría Continental sobre Educación* y OLPED, marzo de 2004. (referido en pág.173 del texto de Feldfeber y Saforcada).

teniendo un costo mucho menor al del aumento de los salarios docentes, la reducción del tamaño de los salones de clase o la reforma de la formación docente. (Wolff, 2007:8).

Esta referencia nos coloca sin pudor ante una expresión más de la razón económica del contexto de las Reformas educativas, pero encontramos a su vez, razones de índole menos explícita relativas al control, al control social y político travestido de control técnico, que pareciera pretender con estas evaluaciones otra *forma* de credibilidad y legitimación social a estas políticas de Reforma. Veamos cómo se produce la incidencia de las evaluaciones nacionales en cuanto a este aspecto.

Las evaluaciones nacionales pueden ser caracterizadas como evaluaciones de tipo *descendente* por cuanto la iniciativa se ubica en la administración central e involucra el poder de **definir** la *calidad* y *valor* de los conocimientos del sistema educativo, estableciendo una *censura de significado*, es decir, una restricción excluyente de otros significados, con valor de única verdad legítima⁶; también llegan a establecer la calidad y el valor de quienes han sido educados y de quienes los han educado.

Es así entonces, que los resultados obtenidos en estas macroevaluaciones, podrán ser explicados, **no** por la gestión político-administrativa, sino por la eficacia y eficiencia de cada institución educativa, o por la deficiente organización de las escuelas y de la 'gestión' áulica. Resultados que son interpretados desde pautas y criterios externos a esa concreta realidad. De hecho, las escuelas pasan a ser unidades de producción de esos productos-resultados donde en las subunidades aulas son 'operarios' los docentes quienes en vez de enseñar, 'gestionan aprendizajes'. Ya es casi de sentido común advertir aquí la descentralización también de responsabilidades mientras centralizado es el pretendido control de resultados del sistema educativo⁷, denominado evaluación. Pero este control es alcanzado en mayor o menor medida según sean las consecuencias de estos resultados y según sea la incidencia en la subjetivación que contribuyen a producir.

⁶ Popkewitz, 1992, 2000.

⁷ Este control se observa de modo notable en Chile donde se busca y se produce la *alineación* contenidos-pruebas nacionales-prácticas áulicas. En Argentina, por el contrario, no se observa tal articulación y la explicación estaría principalmente en que las pruebas son de bajas implicancias ya que sus resultados no producen consecuencias personales ni institucionales. [Radulovich, 2007]

Por otro lado, más allá de las *intenciones*, la *conciencia* o la *voluntad* de algunos agentes gubernamentales, los resultados obtenidos de estas evaluaciones conducen, finalmente, a que toda cualidad de una institución educativa quede *reducida* a una relación coste-beneficio según una ‘función de producción educativa’⁸ a la que se llega a través de complejos algoritmos que nada tienen que ver con desarrollos curriculares contextualizados ni por tanto con los procesos educativos reales⁹. No importa cuántas ni cuáles razones técnicas, estadísticas o pedagógico-didácticas se aduzcan, ése parece ser el destino de tales ‘resultados educativos’ obtenidos a través de las evaluaciones nacionales. No es casual que en Costa Rica se denomine ‘Control de Calidad’ la División del Ministerio de Educación que se ocupa de elaborar, aplicar y procesar la información de las pruebas estandarizadas...

B.

Con algunas diferencias morfológicas, el tipo de prueba de las evaluaciones nacionales en los países de la Región, presenta semejanzas estructurales y homologías funcionales en los ámbitos de pre, inter y post aplicación. Abreviadamente y sin exhaustividad pueden mencionarse, al menos:

Pre

- a. La base del contenido-conocimiento que los equipos técnicos seleccionan como relevante son los *currícula* establecidos como básicos a nivel nacional. En algunos países, intervienen también organismos subnacionales.
- b. Las 'competencias' y operaciones cognitivas se combinan con los contenidos a través de Tablas de Especificaciones en que se cruzan ambos componentes y hasta pueden especificar cuántos ítems de la prueba comprende cada combinatoria. Algunos países dan a conocer estas Tablas a las instituciones con variable antelación a la aplicación de las pruebas.
- c. Son pruebas estandarizadas con opciones para *selección* de respuesta, con muy pocos *ítems*¹⁰ abiertos para *producción* de respuesta. Varían los porcentajes de ítems de

⁸ Mancebón Terrubia, 1999:114 a 123.

⁹ Relación que vincula los *inputs* con los *outputs* de la institución educativa en términos de eficiencia. Se señala la necesidad de cautela en este punto dado que son intangibles tanto el 'producto educativo' como la 'unidad producida'. [Mancebón Terrubia, 1999:116]

¹⁰ Cada una de las actividades de la prueba. Plantean preguntas o cuestiones cuya resolución ha de ser seleccionada de una lista de respuestas, cuatro o cinco usualmente

producción de respuesta de 1% a 3% del total de la prueba, según los países. Los ítems abiertos presentan menor fiabilidad intra e interjueces y mayores costos en el análisis, lo que lleva a reducir tanto la cantidad de estos ítems, aunque sean los más significativos.

- d. El nivel de complejidad cognitiva de los ítems de actividad *no alcanza a las operaciones de orden superior*.

Inter

- e. La aplicación de las pruebas puede ser muestral, censal o ambas, según factibilidad técnica y presupuestaria y según algún criterio político.
- f. Es realizada por agentes de la institución educativa con supervisión ministerial o por aplicadores externos.
- g. En todos los casos existen recaudos de seguridad para preservar el secreto sobre la prueba antes y durante su aplicación... con éxito relativo en cuanto al rigor de este criterio, según los países.

Post

- h. Las respuestas correctas e incorrectas son 'barridas' y cuantificadas por *lector óptico* en los ítems de selección de respuesta, no así en los ítems abiertos o de producción en que las respuestas producidas requieren jueces cuidadosamente seleccionados.
- i. Los *análisis* de calidad de respuesta son realizados por equipos técnicos y las interpretaciones son *formales* pues suponen una única realización cognitiva posible, generalizable y válida para todo contexto socio-cultural y para todo proceso de desarrollo curricular dentro del universo estipulado.
- j. La información de los resultados se difunden a la prensa y/o se diseminan a las instituciones (directivos, docentes) y padres, presentando diferencias en el tipo de información proporcionada según el momento de la evolución histórica del sistema de evaluación de cada país y según los acuerdos de reuniones 'técnicas' en la Región. Pero en todos los casos la información se supone y se acepta como válida técnicamente por parte de los distintos tipos de destinatarios.

C.

Además, hemos observado que en general, los trabajos académicos elaborados para analizar políticas, situaciones educativas o sociales sean de orientación crítica o de orientación

justificatoria hacia las políticas neoliberales, toman a estas expresiones de resultados incuestionablemente como válidos... ¡la *forma* de presentación otorga ilusión de rigurosidad a su *contenido*!, configurando como real una representación sesgada de parcialidades y reducciones de la realidad educativa de referencia. Es decir, el supuesto es que esos datos son certeros o, al menos, confiables.

Sucede que la difusión y diseminación de los resultados de las pruebas distribuyen junto a esta información, un conjunto de *supuestos* implícitos para el público en general, pero también para docentes, directivos y hasta para académicos de diferentes áreas disciplinares.

De tales supuestos, entre otros, podemos mencionar los siguientes:

*Las pruebas estandarizadas de las evaluaciones nacionales proporcionan información pertinente y relevante sobre el nivel de *calidad* que alcanza la educación de una población dada. Este supuesto también alcanza a las pruebas internacionales.

*Las pruebas se elaboran correctamente, toman un campo de contenidos de referencia indiscutible del cual seleccionan aspectos igualmente indiscutibles tanto a nivel de conocimiento como a nivel de operaciones cognitivas. Este supuesto funda lo indiscutible en la *experticia técnica* de quienes las han concebido y elaborado y es compartido respecto de las evaluaciones **internacionales**.

*Los porcentajes o las puntuaciones o cualquiera otro código numérico o gráfico que se presentan como resultados resultan un **signo** válido que *representa* lo aprendido en el sistema educativo y la presencia o ausencia de progresos de los alumnos.

D.

¿Qué podemos decir de estos supuestos? En primer lugar, la noción de calidad no contiene unívoco significado en los distintos países, lo que queda evidenciado en las intensas discusiones en reuniones internacionales -para acordar pautas para las evaluaciones nacionales en Latinoamérica-, según lo reconocen funcionarios de dos ministerios latinoamericanos consultados¹¹. A pesar de ello, **sí** hay acuerdos sobre lo que operativamente se entenderá como indicador de calidad. Es más, estos acuerdos, llegan inclusive hasta los denominados *estándares de currículum* o niveles de logro más

¹¹ Ver Radulovich, 2007.

explícitos¹². Porque el recurso discursivo al significante *calidad* funciona **articulando sentidos** aunque él mismo sea 'flotante' en relación con el significado. Articula las políticas educativas con las expectativas sociales de atención a las 'demandas de un mundo globalizado' en relación con las *competencias*, por ejemplo, produciendo y transfiriendo *valor* a las políticas de gobierno más concretas. En el caso que nos ocupa, aparece legitimando, además, la constelación de resultados educativos y sus justificaciones técnicas. Y, como si esto fuera poco, esta información hasta se ha llegado a tomar como índice para evaluar políticas públicas¹³.

En segundo lugar, se han realizado y se siguen realizando concienzudos estudios que muestran las limitaciones de nivel técnico que presenta la interpretación de los resultados -tanto de las pruebas nacionales como de las internacionales-, como el estudio realizado por Richard Wolfe de la Universidad de Toronto¹⁴. Y... estas advertencias se extienden, a la elaboración, aplicación y usos tanto políticos como a nivel áulico de los resultados de las pruebas. Capitalizando su experiencia de más de quince años en diversos programas de evaluaciones nacionales y regionales, el investigador señala con precisión algunos problemas de estas pruebas que caracteriza como técnicos. Mencionaremos los centrales:

a. La medición y comparación de las pruebas internacionales resultan sesgadas para los países latinoamericanos que se encuentran por lo general en la parte más baja de la escala de logros. Esto sucede porque *...el tratamiento psicométrico y estadístico en esa parte de la escala es particularmente inexacta debido al número pequeño de ítems relativamente sencillos que suelen incluirse en estas pruebas, diseñadas con más altos niveles de expectativas de aprendizaje que las que contienen los currículos de países en desarrollo... existen diferencias entre los marcos de contenidos de las pruebas regionales e internacionales y los currículos nacionales.* [Wolfe, 2007:2]

¹² Ver Documento de PREAL (2006): *Estándares en evaluación*. Implicancias para su aplicación en América Latina.

¹³ Tulic, en Tenti Fanfani [coord.] 2002 *El rendimiento escolar en la Argentina*. Bs. As., Losada; Radulovich, 2004.

¹⁴ Ontario Institute for Studies in Education of the University of Toronto, Richard Wolfe, 2007.

b. Dificultades en las inferencias obtenidas a través de las pruebas pues *...los datos recogidos de manera transversal o en un momento de tiempo [sincrónica]... proporcionan una base débil para analizar los efectos de programas escolares y prácticas [por lo que] ... pueden estar sesgados como para reflejar de manera inversa el fenómeno realmente subyacente.* [ídem].

c. Se produce el "error de especificación"... causado por la *...omisión de variables importantes en el análisis, tales como la oportunidad de aprender [que] ocasiona que otras variables tales como el nivel socioeconómico absorban estadísticamente sus efectos. Los resultados pueden llevar a grandes equivocaciones en la comprensión de los efectos de factores sociales y de las reformas educacionales ...*, por ejemplo.

d. Se efectúan sobreinterpretaciones de correlaciones posiblemente espúreas *... de estudios transversales en un momento del tiempo entre logros estudiantiles y prácticas pedagógicas porque no es posible determinar si fueron tales prácticas las que determinaron dichos logros ... [por lo que] atribuir el conocimiento de los estudiantes a la escuela y los docentes en ese año particular resulta falaz...* (ídem, pág.5)

e. La validez de ítems y pruebas en relación con los contenidos curriculares no está garantizada pues son escasas las situaciones en que se proyecten de modo preciso en las pruebas. Los ítems de opción múltiple tampoco comprenden la suficiente información que permita medir el cumplimiento de los criterios de logro por cuanto *...la selección de una opción.. [no] demuestra por sí misma el proceso [cognitivo] seguido para resolver el problema... la tasa de respuestas correctas está relacionada también con procesos irrelevantes (adivinanza, estrategias para eliminar opciones, etc.).* [ídem, pág.8]

f. La comparabilidad de año en año está afectada pues resultan inválidas o muy imprecisas las comparaciones entre resultados ya que *...las inferencias sobre cambios no son confiables... por errores de equiparación [entre ítems, contenidos y muestras] ...errores de muestreo en ambos momentos... [y] ...diferencias reales.* Entonces, con esta metodología

no es posible medir cambios ni valorar políticas educativas a partir de esos datos. [ídem, pág.10]

g. Los resultados de las pruebas son presentados bajo una *forma* que lleva a una percepción dual: por un lado como demasiado simplificados y, por otro, como difíciles de interpretar. Esto es así porque *...las definiciones de los constructos para las pruebas se realizan a nivel global (lenguaje, matemática, etc.) con referencia a ponderaciones que corresponden a los programas de estudio o a alguna definición de desarrollo intelectual ... este plan asegura una representación adecuada del constructo global en el puntaje total ... se genera una escala para reportar el rendimiento de estudiantes individuales, de distribuciones en grupo, promedios y cambios en el tiempo. Pero los números de la escala son inicialmente arbitrarios, sin significación o interpretación evidente ... no son porcentajes de ningún conjunto de contenidos ... es penoso admitirlo ... nuestra tecnología de evaluación de logros a gran escala no permite medir muchos subconstructos con facilidad [expresión escrita, comprensión lectora, decodificación, resolución de problemas, etc.] ... Esta situación implica que siempre faltará información diagnóstica que permita relacionar el rendimiento diferencial en distintas áreas [matemática, lengua] con factores asociados [capital cultural, ruralidad, pobreza, ambiente de enseñanza, entre otros] ... rendimientos específicos con esfuerzos correspondientes de enseñanza.* [ídem pág. 13 y 14].

A la luz de todo lo expresado hasta aquí, ¿por qué, a pesar de tanta limitación intrínseca de las pruebas, la difusión y diseminación de resultados obtiene algún grado de legitimidad tanto a nivel social general como académico?

Colocando la lupa

Decíamos *ut supra* que analizar el fenómeno de sentidos, razones y efectos de las evaluaciones nacionales parecen proporcionar la posibilidad de una articulación entre la formulación macro y la realización micro de políticas educativas.

Afinando el análisis y retomando el último supuesto implícito transferido al público junto con porcentajes y cifras, decíamos que éstos son mostrados, aparecen, como "un **signo**

válido que *representa* lo aprendido en el sistema educativo y la presencia o ausencia de progresos de los alumnos", y ese signo es prácticamente transformado en un objeto, una cosa con autonomía de su referente real que es el aprendizaje efectivo logrado en unas aulas concretas. Funciona en los discursos oficiales, periodísticos y del público en general como *reificación* de lo representado pues ***el signo se convierte en lo representado, el signo es el aprendizaje***, desplazando el significado o censurando el significado sustantivo. ¿Qué quiere decir o qué adquisición o aprendizaje efectivo revela el 44.49% de incremento en Matemática en primaria al cabo de 8, 10 ó 15 años de aplicación de una Reforma?

Ese mismo signo (el porcentaje) que censura otros significados, *es también criterio de eficacia*, entendida como el valor de las puntuaciones obtenidas en las pruebas. Esto desplaza como criterio de eficacia a los aprendizajes efectivamente obtenidos.

Los resultados porcentuales presentados como ‘cosas’ obtenidas, comparables, combinables, intercambiables, producen un efecto demostración de la bondad de la *formulación* de políticas, en el escenario de su *realización*. Y el criterio de *validación* del signo y de la calidad son las Tablas de Especificaciones, inexcusable *producto experto* como las pruebas mismas y sus interpretaciones.

Cuando las implicancias de los resultados de las pruebas son *altas* como el incremento de asignaciones de fondos, premios salariales a los docentes, y/o acreditaciones a los alumnos, en la institución educativa, en las aulas, el valor asignado, es decir este ‘valor agregado’ a los resultados de las evaluaciones nacionales *opera* como *dirección normativa del proceso de desarrollo curricular* concreto, y finalmente, como *prescripción* de la práctica. Los docentes “espontáneamente” o a través de los directivos y los directivos, “espontáneamente” o a través de los supervisores, reciben el ‘mandato’ de modo explícito o tácito, de ajustar su metodología para que la escuela obtenga mejores porcentajes que la ubiquen o reubiquen favorablemente en la escala o el 'ranking' de escuelas. Sobre todo en Lengua y Matemática.

La reificación o cosificación de los resultados porcentuales producida en el discurso oficial, impacta también la *subjetividad* de los protagonistas de la acción educativa. Si los resultados son magros, los docentes *sienten* aguda *responsabilidad* por los resultados aún en precarias condiciones de trabajo; si las puntuaciones son elevadas, lo que sienten es

satisfacción por la voluntad y el esfuerzo, sostenido a pesar de todo. En el día a día de cada escuela se van construyendo la *convicción* y la *necesidad* de trabajar para obtener mejores puntuaciones. Hay nuevos desplazamientos de sentidos de la práctica, ya no se trata tanto de promover sustantivas comprensiones, conocimiento, reales experiencias de aprendizaje como de administrar el tiempo de enseñanza dentro de una regulación temporal ajustada por los cronogramas de aplicación de estas evaluaciones.

¿Evaluaciones?

Para finalizar, las valuaciones nacionales resultan una aplicación de pruebas para *medir*, no para evaluar, según se señala con claridad en la información y documentos oficiales disponibles de distintos países latinoamericanos, por lo tanto, no se puede afirmar que se trate de una evaluación. Se recoge *información*, desde luego, pero que no es suficiente en cantidad y calidad para formular y sustentar *juicios* acerca de la calidad de la educación. Lo que sí puede, es proporcionar una ‘fotografía’ de cómo los alumnos de distintos lugares de cada país responden a la *norma* establecida en la Tabla de Especificaciones para cada área de contenido, y que son elaboradas centralizadamente en función de acuerdos supranacionales..

Si admitimos a la evaluación como *actitud filosófica* (que busca significados y sentidos) y no sólo como *dispositivo*¹⁵ (que busca sólo la coherencia) ampliaremos sus posibilidades educativas y estaremos en condiciones de colocar como prevalente su condición de *interrogación* acerca de los valores (sean morales, estéticos, políticos, existenciales...), podremos concebirla y vivenciarla como actitud y *marcha crítica* de la experiencia cotidiana. En otras palabras, en la institución educativa y en el aula en relación con la evaluación, se trataría de decidir qué hacer cambiando el orden de nuestras preguntas: discutir, ser conscientes de y acordar *por qué, para qué y para quién* evaluar, **primero** y ver luego el *cómo* coherente con las respuestas que hayamos podido alcanzar para esas preguntas primeras.

Desde ese ángulo de cambio de prioridades en las preguntas, evaluar es -sobre todo-, una proceso bajo el signo de lo inconcluso y referido al orden viviente, al tiempo existencial,

¹⁵ Tomamos aquí el enfoque de Ardoino y Berger, 1989 y Broadfoot, 1984, 2002.

por lo que no puede ser encerrado en un cronómetro o congelado en un indicador. En cambio, lo que habitualmente es denominado ‘evaluación’ en realidad son acciones de **control**. El control en tanto sistema, dispositivo y metodología constituye un conjunto de procedimientos para determinar hasta dónde los objetos o fenómenos responden, se conforman o son idénticos a la norma, al modelo o a la dimensión que se toma como referencia de comparación. Este modelo de referencia es *exterior* y *anterior* (lógica y cronológicamente) a la operación de control misma. Aquí tiene peso fundamental la noción de *corrección*.

A nivel industrial, de organización sistémica de todo orden, el control funciona como control *técnico*, necesario para optimizar la acción del mecanismo, la eficacia práctica del sistema, en un curso previsible y preciso.

En cambio, cuando se traslada esta concepción de control -sustentada en la noción de corrección- a las relaciones sociales aparece una superposición de *control social* y control técnico, operando un efecto de ejercicio de poder sobre otros bajo la supuesta y necesaria comprobación y/o verificación técnica. En educación, esta indiferenciación entre control técnico y control social se puede advertir en la forma en que la así llamada evaluación nacional de los sistemas, va ‘modelando’ progresivamente pensamiento, acción y sentir de los agentes educativos a lo largo de tiempos y espacios cada vez más extendidos y profundos dedicados a este tipo de actividades ‘evaluativas’ pudiendo llegar a ejercer un poder de determinación fuerte. Y así, en el caso de las situaciones denominadas de evaluación, éstas tienden a *modelizar* el pensamiento y la acción ajustándolos a una normatividad *lógica* de identidad, conformidad, compatibilidad y coherencia y a una normatividad *moral* a través de la noción de **corrección** a través de la cual se pasa de lo verdadero al bien o a lo bueno... según relativos, muy relativos e interesados criterios supuestamente objetivos.

Entonces, en esos términos, en las concretas personas de la institución educativa (directivos, docentes, alumnos) en el convencimiento del *buen* hacer para obtener una *buena* evaluación, se produce la **sujeción** a la supuesta eficacia práctica de la optimización de la acción cuando en realidad lo que se produce es la *eficacia simbólica* del ejercicio del poder por la articulación entre constatar y sancionar en el caso de no encontrar, el evaluador, el ajuste esperado.

Apreciamos, a partir de todo lo expuesto, la operación oculta y esencial de las funciones políticas **primarias**¹⁶ de los sistemas nacionales de evaluación, es decir la gestión productivista del sistema, el control administrativo y la selección de individuos, a la vez que que operan como *...estrategia de legitimación compensatoria, proporcionando credibilidad tranquilizadora confiriendo al proceso político el valor agregado de la aceptabilidad científica*”(Wieller)¹⁷.

Y a nivel de los sujetos tanto docentes como directivos, producen al menos dos efectos no visibles pero presentes de *legitimación 'resignada' por culpabilización y prescripción de la práctica por normatización*.

No estamos planteando que los docentes carezcan de responsabilidad social en relación con su función social, hay una responsabilidad social del trabajo docente con respecto a sus alumnos, con respecto a la comunidad de la que forma parte, con respecto a sus colegas y con respecto a sí mismo también. Responsabilidad en cuanto a promover *aprendizajes sustantivos* no sólo en relación con las áreas de conocimiento, sino también en cuanto a los *valores intrínsecos* que hace educativa cada actividad que proponga. Y por sobre todo, en cuanto a pensar, analizar las realidades, desbrozando lo superficial, aparente o ficcional para *recuperar* así los sentidos desplazados y, entonces, poder *educar* a nuestros alumnos. Resultaría más factible en tal caso, neutralizar la posibilidad de oficiar de operarios de esta lógica productivista. Responsabilidad social de los docentes sí, pero **no** responsabilidad sobre aquello que los -o nos- sobredetermina desde lejos.

¹⁶ Angulo Rasco, 1995:203-207.

¹⁷ Wieller (1992); y también en Kvale (1992); House (1992, 1998); Carnoy (1999); Mancebón (1999); Kogan (1992, 1996); Lundgren (1992, 1996); Tiller (1992), Angulo Rasco (1995, 1999), sin agotar la enumeración.

REFERENCIAS

- ANGULO RASCO, J.F. “La evaluación al sistema educativo: algunas respuestas críticas al por qué y al cómo” en *Volver a pensar la educación. Vol II: Prácticas y discursos educativos* (Congreso Internacional de Didáctica). (1995) Madrid, Morata.
- ARDOINO, Jacques y BERGER, Guy (1989) *De una evaluación en migajas a una evaluación en actos. El caso de las Universidades*. Láigle Matriz, ANDSHA
- BROADFOOT, Patricia “Evaluación e intuición” en ATKINSON, T. – CLAXTON, G. [Edits.] (2002) *El profesor intuitivo* Barcelona, Octaedro.
- BROADFOOT, Patricia (1984) ‘From public examinations to profile assessment’: the French experience’ en Broadfoot, P. (Ed.) *Selection, Certification and Control*, London, Falmer Press.
- BRACEY, G.W. (2002) *Sobre pruebas y mediciones de aprendizaje: un breviario para la “alfabetización evaluativa”*. Lima, PREAL y GRADE
- CALIDAD DE LA EDUCACIÓN (2003) *Evaluación de aprendizajes para una educación de calidad*. Ministerio de Educación de Chile.
- DUSSEL, I., TIRAMONTI, G. y BIRGIN, A. ‘Hacia una nueva cartografía de la reforma curricular. Reflexiones a partir de la descentralización educativa argentina’ en *Revista de Estudios de Currículum* Volumen 1, N° 2, abril de 1998 (pp. 105-161)
- FELDFEBER, M. y SAFORCADA, F. (2005) *La educación en las Cumbres de las Américas*. Bs. As. Miño y Dávila.
- FERRER, G. (2006) *Sistemas de evaluación de aprendizajes en América Latina. Balances y desafíos*. Publicación de PREAL.
- HOUSE, Ernest (1994) *Evaluación, ética y poder* Madrid, Morata.
- McMEEKIN, R.W. ‘Hacia una comprensión de la *accountability* educativa y cómo puede aplicarse en los países de América Latina’ en CORVALÁN, J. y McMEEN, R.W. [edit.] (2006) ‘*Accountability*’ *educacional: posibilidades y desafíos para América Latina a partir de la experiencia internacional*. Santiago de Chile, Preal-Cide.
- MANCEBÓN TERRUBIA, M^aJ. ‘La función de producción educativa: algunas conclusiones de interés en la especificación de los modelos de evaluación de la

- eficiencia productiva de los centros escolares' en REVISTA DE EDUCACIÓN N° 318. Madrid, Ministerio de Educación y Ciencia.
- POPKEWITZ, Th. (1992) "Algunos problemas y problemáticas en la producción de la evaluación" en *Revista de Educación* N° 299 Madrid, MEC.
- POPKEWITZ, Th (2000) "The denial of change in educational change: Systems of ideas en the construction of national policy of evaluation" *Educational Resercher*, Vol. 29 N°1 ER Online <http://www.aera.net/pubs/er/arts/29-01/popkew01.htm>
- RADULOVICH, S. (2003) "Evaluaciones nacionales y homogeneización cultural: ¿qué sucede con las evaluaciones de Lengua?" en Fernández, A. y Rodas, J. (Com.) *Historia y sociolingüística del español en el noroeste argentino*. Salta, publicación de la Universidad Nacional de Salta.
- RADULOVICH, S. (2004) *Evaluaciones de sistemas educativos en la Región Latinamericana, ¿para qué?*. Ponencia al 2º Congreso Internacional de Educación, Universidad Nacional del Litoral, Santa Fe.
- RADULOVICH (2007) *Algunas consideraciones analíticas sobre las bajas y altas implicancias de las evaluaciones nacionales en Argentina y Chile*. Ponencia en las I Jornadas de Investigación Educativa, Mendoza, mayo 2007.
- TULIC, L. 'La evaluación de las políticas públicas' en Tenti Fanfani, E. [comp.] (2002) *El rendimiento escolar en la Argentina Análisis de resultados y factores*. Buenos Aires, Losada.
- WOLFF, L. (1998) *La evaluaciones educacionales en América Latina: avance actual y futuros desafíos*. Boletín N° 11 de PREAL (Programa de Promoción de la Reforma Educativa en América Latina y El Caribe).
- WOLFF, L. (2007) *Los costos de las evaluaciones de aprendizaje en América Latina*. Documento N° 38 PREAL.
- WOLFE, R. (2007) *Cuestiones técnicas que condicionan las interpretaciones de los datos generados por las evaluaciones de logros de aprendizaje escolar en América Latina*. PREAL, Grupo de Trabajo sobre Estándares y Evaluación.
www.preal.org/grupo2.asp?ld_Grupo=3