

Las evaluaciones y la transición hacia la Universidad crece el optimismo de los empresarios para 2005

Enrique G. Costa Lieste, Decano Facultad de Ciencias Empresariales – UCES

Urgencia y problemas de las evaluaciones nacionales

Desde las décadas de 1960-70, debido al exceso de la demanda escolar sobre la oferta, diferentes países, industrializados o no, comenzaron a utilizar evaluaciones nacionales para ordenar el acceso a los niveles superiores de educación, y esa necesidad se reforzó con la aceleración de la realidad económica de la década de 1990 y principios del siglo XXI, que no permite mayores dilaciones para la adecuación de los alumnos a la vida real.

Comenta Jorge Hernández Uralde del CENEVAL de México:

Sólo los demagogos y los ingenuos pueden pretender que la vida escolar de toda persona comience en un jardín de niños y llegue indefectiblemente a la educación superior. En México todavía hay personas que piensan así y algunas incluso lo hacen de buena fe (Ministerio de Cultura y Educación de la Argentina, "Sistemas de Admisión a la Universidad": 89)

Características más comunes de los regímenes de evaluación

La norma más común que rige la admisión a la universidad en países como Inglaterra, Francia, España, Italia, Brasil, Colombia, Chile y México es la exigencia de un examen a la salida del nivel secundario sobre la base de los estudios realizados. Los certificados de estudios son una exigencia general, aunque varía la importancia que se les atribuye. En algunos países los promedios de los estudios cursados son combinados con los resultados de los exámenes de evaluación y ponderados con diferentes pesos. En Alemania, por ejemplo, las calificaciones del secundario tienen una ponderación del 70% en la nota final mientras que en España es del 50%. El país más exigente de los mencionados es Colombia, donde se toman en cuenta las calificaciones de los certificados finales, más las notas de los exámenes obligatorios a la salida del secundario, más las notas de un examen de admisión a la universidad.

En los EE.UU. no existe un régimen oficial de evaluaciones pero está generalizado el uso de un sistema privado que las universidades suelen aceptar como condición de ingreso. Las principales críticas que recibe dicho sistema hacen referencia a la metodología, y específicamente a la técnica de las *multiple choice*. Dado que esa técnica se emplea también en las evaluaciones nacionales de carácter público de los países que las han establecido en forma obligatoria, consideré necesario profundizar más en detalle el sistema de los tests de los EE.UU., como un antecedente para tener en cuenta en el programa vigente de evaluación escolar en la Argentina.

EE.UU.: los IQ y los SAT

SAT es la sigla que identifica en los EE.UU. al Test de Aptitud Escolar (*Scholar Attitudinal Test*) y, desde 1993, al Test de Evaluación Escolar (*Scholar Assessment*)

Test). En estas páginas solamente me referiré al SAT I porque según sus autores intenta medir “el razonamiento”, además de la importancia que tiene el hecho de ser tomado por unos dos millones de estudiantes anuales (*The College Board: Ten Real SATs 1997: 2*) frente a solamente 230.000 del SAT II (*The College Board: Real SAT II 1998:v*).

Los SAT son producidos por el ETS (*Educational Testing Service*), una empresa privada fundada en 1947 que a su vez depende del *College Entrance Examination Board* (más conocida como *Board*) establecida en 1900. Esta es un tipo de fundación sin propósito de lucro que cada año nombra a un *board* de expertos para diseñar los SAT. En un comienzo los tests consistían en ensayos bastante largos pero, debido a la “fascinación” que provocaron los tests de inteligencia, se incluyeron en 1926 las *multiple choice* hasta que en 1942 los ensayos se abandonaron definitivamente a favor de las opciones múltiples. La razón dada por el *Board* fue que, con motivo de la II GM, fue muy difícil conseguir *graders* (“liquidadores” de puntuaciones) y se debió recurrir a las opciones múltiples. Sin embargo, al final de la guerra no se volvió al sistema de ensayos: “los viejos exámenes eran simplemente muy difíciles de producir y muy costosos para calificar” (Owen-Doerr: 27).

Las aptitudes generales

En los primeros años del siglo XX los maestros de los EE.UU. jugaban a adivinar qué texto de Virgilio o escena de Shakespeare elegiría el *Board* para el próximo año. Para eliminar este problema, en 1916 el *Board* dio el primer paso *para apartarse de preguntas que pudieran identificarse con currículos vigentes*, haciendo más “comprehensivos” o amplios los contenidos de las preguntas. A medida que se fueron utilizando los tests del ETS se hizo difícil saber qué era lo que estaban midiendo.

El *Board* consideró que deberían medirse ciertas cualidades importantes vinculadas con algo vagamente definido como “vigor intelectual” o poder mental, pero ciertamente “el nuevo énfasis no estaba puesto en el conocimiento”. El *Board* pensaba que así se podría medir la “aptitud primaria que surgía más allá de una niebla de hechos y cantidades”. Estas aptitudes primarias postuladas por los miembros del Board recuerdan el intento de medir una supuesta inteligencia general mediante los IQ. Quienes se opusieron a este cambio adujeron que “los nuevos tests premiarían a una sabiduría superficial a expensas de lo académico” (Owen y Doerr: 92).

El entrenamiento para los SAT

Como la intención era medir una capacidad mental general, la ETS sostenía que los resultados de los tests no podrían, por lo tanto, ser factibles de mejorar con tutores o maestros privados. Hoy, sin embargo, todos saben que esa ayuda eleva algunos puntos los resultados y esa es la razón de la existencia de importantes “academias” que enseñan cómo “batir” a los tests. La red nacional Kaplan ha testeado a más de tres millones de estudiantes (xiv) y Gary Gruber se enriqueció con la venta de sus manuales (más de cinco millones vendidos: “el más efectivo y completo libro de SAT en el mercado...¡Estrategias probadas!”).

Cuando se puso en evidencia que los tests podían realmente enseñarse, la palabra “Aptitud” de la sigla SAT comenzó a molestar al *Board*. Las nuevas dificultades se centraron en la confusión que se originaba con el empleo de la palabra “habilidades”. El

vocablo habilidad se confundía con aptitud y logro. Nada menos que el *National Research Council* tuvo que producir un informe acerca del significado de las pruebas de habilidades. La conclusión fue que las pruebas de aptitudes y de logros son ambos tests de habilidades. Existe “muy poca diferencia entre los términos porque la habilidad es siempre una combinación de aptitud y logro”. La habilidad se suele definir como “el límite superior de lo que una persona puede hacer ahora, pero no debe confundirse con potencial o capacidad” (194). La palabra Aptitud en la sigla SAT se abandonó en 1993, cambiándose por “*Assessment*” (Evaluación). En esta forma no se cambió la sigla, pero “actualmente no significa nada, es solamente un nombre” (Owen-Doerr: 193).

Los problemas de las opciones múltiples

Toda la discusión acerca de las críticas que se le han hecho a las *multiple choice* se traslada a los SAT, dado que estos son diseñados en base a opciones múltiples: La más antigua controversia que involucra a la ETS se refiere a lo correcto de los tests de opciones múltiples. Sin esos tests, que pueden ser calificados por máquinas, el testeo de millones de personas cada año sería imposible y la ETS estaría en bancarrota. Más importante aun, la amplia aceptación del testeo standarizado depende de la percepción de que es científicamente neutral y objetivo, lo que solamente un test con respuestas “correctas” o “equivocadas” puede lograr (Owen – Doerr: 27).

Los diseñadores de las *multiple choice* tienen forzosamente que producir una discriminación entre los alumnos, si no las preguntas no sirven. Proceden entonces a redactar preguntas *desorientadoras* : “no siempre lo hacen intencionalmente, pero siempre lo hacen, en parte porque es difícil no hacerlo”, o a “utilizar *distractores* con palabras que los alumnos nunca han visto, para que no puedan imaginar la respuesta que se espera” (45).

Doerr, al concluir la actualización de 1999 del libro de Owen de 1985, declara estar convencida de que las pruebas normalizadas hacen poco bien y mucho mal. Más aun, existe una organización no gubernamental llamada *FairTest* cuyo propósito es la eliminación lisa y llana de cualquier tipo de test standard (Doerr en Owen-Doerr: xii).

La verdadera selectividad de las universidades

Al contrario de lo que puede ser un estereotipo para el público del exterior en lo que respecta a los criterios de selectividad de las universidades norteamericanas, dicen Owen y Doerr que, comparado con otros países, todo el sistema de testeo, educación y certificación parece no importarles a las instituciones educativas de los EE.UU.:

Unas pocas docenas de entre varios miles de *colleges* y universidades enfrentan una competencia seria para las admisiones. Solamente 25 universidades grandes y 60 pequeños *colleges* tienen porcentajes de admisión menores al 50% de los solicitantes. La gran mayoría admite a casi todos los que entregan solicitudes (xv).

Si esto es así: ¿Por qué los *colleges* requieren los SAT si no tienen ninguna intención de usarlos? Dos autores que investigaron la conducta del ETS dijeron que:

Llegamos a la sospecha de que la mayor parte de las instituciones que no son selectivas pero que sin embargo requieren los tests lo hacen para mantener el aura de selectividad...No les cuesta nada. Los pagan los chicos (Hartnett y Feldmesser en Owen-Doerr: 230).

Cuando los distritos escolares, a su vez, aprenden que son juzgados por los padres de acuerdo a las puntuaciones de los SAT de sus alumnos, reconvierten todo sus sistemas

educativos y los transforman en centros en gran escala de preparación para los tests (xiv).

La distorsión del aprendizaje

De allí en adelante la modalidad de lectura de los alumnos se distorsiona desde el aprendizaje de conocimientos hacia el aprendizaje de habilidades para aprobar los SAT, que no es lo mismo (Sternberg 1966: 166).

Gruber afirma que “las habilidades para el Pensamiento Crítico se han convertido en la voz de orden en los EE.UU.” (“es la frase que zumba en la nación”) y las define como “habilidades *genéricas* para la forma más efectiva y creativa de resolver un problema o evaluar una situación”. Sin entrar a analizar la corrección del razonamiento de Gruber acerca del razonamiento, o sobre la supuesta ventaja de utilizar el “pensamiento crítico” como metodología de evaluación, dicho autor define así al pensamiento crítico: Antes que nada, creo que la inteligencia puede ser enseñada. La inteligencia, definida simplemente, es la aptitud o habilidad para razonar las cosas. Estoy convencido que Ud. puede aprender a pensar lógicamente y resolverlas mejor y más rápido, particularmente en relación a los problemas de matemática y de lenguaje del SAT. Pero alguien tiene que darle las herramientas. Llamemos estrategias a estas herramientas. Y esto es en lo que consisten las Habilidades de Pensamiento Crítico: estrategias (Gruber 1998: v).

Las instrucciones para lograr esos resultados son:

- Extraiga u observe cualquier cosa curiosa (en el texto de la pregunta).
- Utilice estrategias específicas junto con habilidades básicas (xvi).

Un tipo de preguntas incluidas en los SAT se denomina Analogías Verbales. Como ejemplo de aplicación de las dos reglas de arriba, Gruber (Ph.D.) ofrece la siguiente Analogía Verbal, ejercicio que consiste en elegir un par de palabras que tengan sentido en relación al par que se pone como guía:

Complete la Analogía:

LADRON: ROBO
Hurtador: diamantes
Cárcel: crimen
Secuestrador: rehén
Criminal: captura
Falsificador: adulteración

La manera no-Crítica: relacione las palabras, LADRON, hurtador, ROBO, diamantes, o LADRON, cárcel, ROBO, crimen, etc.

La manera Crítica: Relacione las palabras LADRON y ROBO en una oración: el LADRON comete un ROBO. Ahora ponga cada una de las palabras de las opciones en la misma forma de oración:

El *hurtador* comete un *diamante* - ¡no tiene sentido!
La *cárcel* comete un *crimen* - ¡no tiene sentido!

Así continúa el ejemplo hasta que:

El *falsificador* comete *adulteración* – tiene sentido y es la única opción disponible.

Para el alumno las preguntas se reducen, en realidad, a un examen de vocabulario, porque para resolver los ejercicios se deben conocer los significados de palabras que en inglés no son usuales (aunque las sean traducidas al castellano). Por eso el manual incluye “las 50 habilidades de Pensamiento Crítico de Gruber”, aunque en realidad son 52 (Gruber: xvii). Gruber presenta 5 estrategias *generales* para el pensamiento crítico y 52 específicas que los alumnos pueden estudiar para elevar sus chances de obtener buenas puntuaciones. Además, el Manual ofrece un listado de 3.400 palabras, otras 291 que son las más utilizadas en los nuevos SAT de lenguaje, 366 raíces latinas y griegas, prefijos y sufijos, que se supone deben memorizarse, todo esto para resolver las Analogías Verbales y los ejercicios de Completar Oraciones.

Como resultado:

Las horas que se invertían antes leyendo libros y haciendo verdaderas tareas ahora están dedicadas a cargar la cabeza con vocabulario y aprendiendo a cómo elegir entre las analogías del ETS (Owen-Doer: 258).

Cómo ser sabio con los tests

El *Board*, por su parte, publica un manual denominado “Los 10 SAT reales” (1997) donde asegura que es “la *única* fuente de las verdaderas preguntas de los SAT de *tapa a contratapa*”, en una alusión evidente a Gruber. Al pie de la cubierta se puede leer:

Las *últimas* estrategias para dar el examen, desarrolladas por los constructores de los tests.

Descripciones *detalladas* y sólida práctica para toda clase de preguntas verbales y de matemática.

He aquí algunas de las estrategias recomendadas por el *Board*. Tener en cuenta que siempre se presentan las *múltiple choices* con cinco diferentes resultados para optar, y que cada minuto de las tres horas permitidas es importante:

SI UD. NO SABE LA RESPUESTA CORRECTA, es más fácil comenzar por eliminar las que Ud. sabe que son incorrectas y (*guess* : adivine, acierte) con las que restan (Board: 12, 17).

Estime cuando puede eliminar al menos una pregunta: cada una correcta, 1 punto; en blanco: ningún punto; incorrecta: $\frac{1}{4}$ de punto (12). (Se está aconsejando que es preferible dejar preguntas en blanco).

Todas las preguntas tienen el mismo puntaje. No pierda tiempo. Resuelva las fáciles antes (12).

Los que rinden buenos tests desarrollan un sentido del tiempo para completarlos (15).

(Se supone que los alumnos se hacen hábiles con los tests completando los de prueba

que les ofrecen las academias, los manuales y los mismos cursos preparatorios de las escuelas...o rindiendo los SAT varias veces).

Sazonemos estas “estrategias” con los siguientes comentarios de Owen-Doer:

Virtualmente todos los distritos escolares reconocen ahora el concepto de ser *test-wise* (“sabio con respecto al test”): imaginarse la mentalidad que el ETS aplica cuando escribe los tests (Owen-Doer: xiii).

Al enfatizar la habilidad por sobre el contenido, el SAT manda una señal trágica a los profesores y a los directivos (259).

Los SAT y el nivel de calidad

Unas palabras en cuanto a las tendencias de las puntuaciones de los SAT. Sabemos que en los tests clásicos de IQ se están observando subas en las puntuaciones (el denominado efecto Flynn). En lo que hace a los SAT, por el contrario, lo que se viene observando es una baja en los *scores*. Los SAT comenzaron siendo utilizados por escuelas prestigiosas para consolidar sus famas. Como sus alumnos pertenecían a las altas clases sociales, sus IQ eran más elevados que los alumnos de las escuelas más modestas. Cuando la presión social hizo que cada vez más escuelas se vieran obligadas a entrar en el circuito de los SAT (que no son obligatorios en lo legal pero sí en los hechos), lógicamente los promedios nacionales de las puntuaciones de los SAT comenzaron a disminuir (Sternberg 1996: 83).

Sí existiría una disminución en los conocimientos adquiridos por los alumnos norteamericanos, que se correlacionan con la baja en los EE.UU. del *nivel de los textos* que solicitan los profesores. Los libros de ahora estarían tres niveles por debajo de los que se utilizaban hace treinta años atrás, de acuerdo a una investigación financiada por el Departamento de Educación y que implementó Sally Reis de la Universidad de Connecticut. Lo mismo sucede con los libros de texto de las escuelas primarias. “Como autor de libros de texto –dice Sternberg- encuentro que nunca estoy bajo presión para elevar el nivel de lo que estoy escribiendo. La presión es siempre para bajar el nivel”: Cuando los distritos escolares compran libros –aun cuando los redactan profesores universitarios- ellos “compran fácil”, de manera que sus estudiantes no se queden muy perturbados. Pero éstos no se enfrentan tampoco con algo que les signifique un mayor desafío (84).

Finalmente, al forzar a los alumnos a recurrir a la memoria para luego “regurgitar una gran colección de hechos triviales y desconectados entre sí” hace que el test de aptitud “sea un examen cuya única materia verdadera es el mismo test” (257). Wurman se quejaba en 1989 porque “el extraordinario énfasis en la memoria a corto plazo a expensas de la comprensión a largo plazo” estaba extendido en todo el sistema educacional de los EE.UU. (Cavanaugh: 113). Este sería el resultado de la “virtual conspiración” de los IQ contra la evaluación de la comprensión a la que se refiere Gardner, sobre todo cuando se trata de aplicar enfoques múltiples que apelan explícitamente a amplios rangos de inteligencias, capacidades e intereses, y que hacen de la evaluación una tarea verdaderamente compleja (1999a: 176). Aunque la medición mediante IQ parece atravesar en la actualidad su gran crisis, los tests de inteligencia ganaron la guerra porque si bien fueron reemplazados por los SAT en los EE.UU., estos no serían mas que "clones de los IQ" (18). Al cambiar el énfasis desde el aprendizaje

hacia las puntuaciones de los SAT, “ETS ha contribuido en forma importante a la verdadera declinación en la calidad de la educación en los EE.UU.” (Owen-Doerr: 258)

¿Qué se puede hacer contra los SAT?

Si bien el proceso de una evaluación nacional que implica a millones de alumnos no se equipara a un examen a treinta alumnos en un aula, “las características de los SAT, que los hacen rápidos y de fácil administración, son las mismas que hacen que sus contenidos sean tan superficiales, injustos y sus efectos sobre las escuelas tan perniciosos” (269). Parece ser que ante esto no hay compromiso posible, salvo el hacer una docena de preguntas vinculadas con el currículo para ser contestadas una mañana de un sábado hacia el final de la secundaria. ¿No sería la mejor forma de juzgar la preparación de un estudiante que pretende cursar los estudios superiores?” (267). La propuesta de la cita coincide con la metodología que se aplicaba para tomar los exámenes de ingreso a las escuelas públicas secundarias en la Argentina de los tiempos mejores. Ahora estamos en las inmediatas puertas de convertir en obligatorias las pruebas finales nacionales para los egresados de las escuelas secundarias. El peligro es que tomemos el camino simbolizado por el ETS, sea privado o estatal. (Ver más adelante la opinión de Gardner sobre este tema, tal como se la expresó al entonces ministro Llach).

El caso del Brasil

Desde 1996, con la ley de Directrices y Bases de la Educación, el Brasil le fijó un papel determinado a la enseñanza media, desvinculándola de su tradicional función de ser una etapa hacia la universidad. Desde 1915 se ha tratado de democratizar la enseñanza superior mediante exámenes de “antesala”, de ahí su denominación de vestibulares.

El INEP es la agencia del gobierno responsable de los Exámenes Nacionales de Enseñanza Media (ENEM), y de los exámenes para los egresados de la educación superior conocidos con el nombre de “la Aprobación” (*O Prova*). Según la explicación del INEP, las tendencias internacionales acentúan la importancia de la formación general en la educación básica no solamente para la continuidad de la vida académica sino también para una actuación autónoma del sujeto en la vida social, destacando su inserción en el mercado del trabajo, que se torna más y más competitivo...(Los contenidos tradicionales deben estar asociados) al desarrollo de estructuras capaces de *operacionalizarlos en el enfrentamiento con los problemas* que surgen en la realidad social, cada vez más compleja, y en una dinámica de tiempo progresivamente acelerada. (Ver dirección *web* en Bibliografía). (Énfasis agregado).

El examen final de la enseñanza media

Los exámenes del ENEM son *voluntarios* y sus resultados, confidenciales, sirven como base para los criterios de admisión que fijen las universidades mediante los cortes en las puntuaciones de acuerdo a la cantidad de plazas. El alumno que participa del ENEM “puede avalar *su capacidad para utilizar los conocimientos adquiridos en la escuela para la resolución de los problemas* que se le presentan en el día a día, interpretando datos e informaciones”.

Los exámenes de “salida” para la enseñanza media del ENEM buscan atacar las raíces

del síndrome del alumno inerte en sus dos bases, el conocimiento inerte y la desmotivación, mediante la verificación de si el conocimiento así construido *puede ser efectivizado* por el alumno por medio de la demostración de su autonomía de juicio y de acción, de actitudes, valores y procedimientos frente a *situaciones-problema que se aproximen lo máximo posible a las condiciones reales* de la convivencia social y del trabajo individual y colectivo. (Énfasis agregado).

La representación de situaciones problemáticas que se aproximen “lo máximo posible” a la realidad, incluyendo principalmente el trabajo, se realiza mediante la presentación de situaciones concretas (casos) mediante los cuales los alumnos “deben organizar las informaciones y los conocimientos disponibles”. Sin embargo, la forma en que esto se hace es la parte criticable de la tarea del INEP, como se tratará de demostrar.

Las críticas a los SAT y los casos de examen en Brasil

Los casos que integran las evaluaciones brasileñas sirven de ejemplo exacto para las argumentaciones que se han leído en cuanto al empleo de los SAT en los EE.UU. Se transcribe a continuación uno de los casos correspondientes a la prueba de los cursos de grado de Administración, para analizar luego las implicancias de la puntuación en base a opciones múltiples en el tema específico del método de casos.

El presidente de una gran tienda por departamentos quiere aumentar las ventas y decide hacer una liquidación de calzados, porque ese departamento está situado al fondo del segundo piso de la tienda. El presidente piensa que los clientes comprarán otros productos durante el trayecto hasta el fondo del segundo piso. Como el departamento de calzados es el más distante desde la entrada del local, solamente una liquidación fantástica atraería a los clientes al segundo piso. El presidente decide, entonces, hacer una liquidación del tipo “Pague 1 y lleve 2”. Entretanto, el gerente del departamento tiene como meta vender el mayor número posible de zapatos, porque sus vendedores están a comisión y dependen del valor total de las ventas. Cuando el presidente anunció que el departamento de calzados tendría una liquidación del tipo “pague 1 y lleve 2”, el gerente se enfureció.

Podemos identificar la fuente de este conflicto organizacional en la:

- a) Competencia por los recursos entre sectores o departamentos.*
- b) Falla en la comunicación y en la interpretación de la información.*
- c) Diferencia entre las metas de los grupos (departamentos o unidades).*
- d) Divergencia sobre los criterios de departamentalización.*
- e) Incongruencia de la estructura de la empresa*

El alumno tiene que limitarse a marcar una de las opciones. Si no supiera nada de Administración, podría elegir una respuesta al azar y tendría un 20% de chance de acertar con la respuesta correcta; por eso, el examen procura incluir varios tipos semejantes de preguntas para eliminar aquel efecto. Si el alumno sabe algo de Administración, la respuesta no le exige precisamente un gran esfuerzo. Y más aun, si una persona no ha estudiado Administración ni tiene experiencia en empresas (por ejemplo: un filósofo, salvo el caso de página 50), podrá satisfacer la pregunta empleando solamente su sentido común, que por definición no forma parte del currículo universitario.

Efectos de las opciones múltiples en Japón

El efecto de la técnica de las opciones múltiples también se evidencia en Japón, al que tantas veces se lo pone como modelo. Los tests nacionales de evaluación comprimen en tal forma el desarrollo de los temas, que obligan a los alumnos a memorizar entre diez a veinte rubros por página sin darles ningún material textual para 'masticar' y ninguna base real para la discusión en clase o para la especulación individual. Las escuelas griegas de pensamiento, por ejemplo, están encapsuladas típicamente en una sola oración. Como a los alumnos japoneses lo único que les interesa son los resultados del test, no quieren perder tiempo aprendiendo temas que no van a estar incluidos en los mismos (Owen-Doerr: 256).

Las trágicas narraciones acerca de lo que representa para los alumnos del Japón no lograr el puntaje suficiente para poder entrar en la universidad se ilumina ahora con una luz distinta, sobre todo sabiendo que luego de haber hecho el esfuerzo de memorización que implican los tests llega el anticlimax .

La mayoría de los estudiantes japoneses no hacen gran cosa en las universidades a las cuales han dedicado sus vidas para ingresar. “Ud. se despierta tarde y juega un montón al mahjongg”. Muchos estudiantes consideran a sus cuatro años de *college* como a una bien merecida vacación –las primeras reales vacaciones de sus vidas (257).

Propuestas en la Argentina

En agosto de 1998 se realizó en Buenos Aires un seminario internacional acerca de los “Sistemas de Admisión a la Universidad” organizado por la Secretaría de Políticas Universitarias del Ministerio de Cultura y Educación. En su presentación, la entonces titular del Ministerio, Susana Decibe, mencionó que la articulación adecuada entre el nivel medio y la universidad era uno de los desafíos que todavía debía enfrentar la educación argentina:

En este sentido, se observa que mientras la mayoría de los países han instaurado diversas prácticas de selección de los estudiantes para el ingreso al nivel superior, en la Argentina se ha mantenido por mucho tiempo la idea del “ingreso irrestricto”, sostenida por distintos sectores universitarios como un verdadero dogma (12).

El Consejo Nacional de Educación Superior del Ministerio aprobó un dictamen en 1994 (un año antes de que Brasil comenzara su nueva experiencia con el INEP), que recomendaba

la instrumentación, al término de la Educación Media o Polimodal, de una prueba *voluntaria* de conocimientos y aptitudes...Ese certificado podría constituir un elemento más de juicio para aquellas instituciones de nivel superior que, en ejercicio de su autonomía, lo incorporaran a sus sistemas de admisión, junto al tradicional certificado analítico de estudios (énfasis agregado).

El CNES recomendaba, además, en conformidad con el artículo 53, inciso e, de la Ley Federal de Educación, implementar “programas *remediales*” para completar la capacitación de quienes no lograsen aprobar dicha prueba en su primer intento. Dicho artículo expresa:

Art. 53. El Poder Ejecutivo nacional, a través del ministerio específico, deberá: [.....]

e) Implementar programas especiales para *garantizar* el ingreso, permanencia y *egreso* de los alumnos/as en todos los ciclos y niveles del sistema educativo nacional, en coordinación con el Consejo Federal de Cultura y Educación. (CNES) (Énfasis agregado).

Si bien el populismo ha dejado su marca en las palabras *garantizar* y *egreso*, el dictamen proponía que los alumnos que no lograsen demostrar una mínima aptitud podrían presentarse nuevamente en años sucesivos, tantas veces como lo desearan (III, 8). Y refiriéndose al artículo 53, el CNES propuso que el PE debería contribuir al financiamiento de una capacitación complementaria para cumplir con los principios de equidad e igualdad de oportunidades. Nada dice, por supuesto, acerca de cómo garantizar el *egreso en todos los niveles*.

Posteriormente, en junio de 1997, el CNES, atento “a los comentarios y sugerencias a que dio lugar dicho dictamen”, y en razón del artículo 5 (incs. f. y g.: igualdad de oportunidades y equidad para el logro de “la mejor calidad posible”) propuso, en una medida que parece muy acertada, “anticipar la puesta en práctica de algunas medidas remediales con carácter tanto correctivo como preventivo” mediante una Prueba Preliminar *obligatoria* al término de la EGB (Educación General Básica) con el propósito de poner en evidencia las insuficiencias o defectos formativos de los alumnos que sirvan de base para el diseño de los programas denominados remediales.

Medir lo que se debe medir

No parecía haber mayor oposición en la Argentina a finales de los 90 para continuar con las próximas etapas del proceso de evaluación educativa comenzado en 1993. Tanto los padres como sus hijos y los docentes aprobaban su utilización en porcentajes que oscilan entre el 89 y el 97%. Incluso una gran mayoría aceptaba que los resultados fueran de conocimiento público (75 al 82%) (Llach et al.: 331).

Si la opinión pública fuera favorable, sería importante perfeccionar las técnicas que buscan medir la comprensión de los conceptos claves de las disciplinas, dentro de una dimensión del largo plazo. Entre dichas técnicas se cuentan las pruebas mediante los ensayos de redacción, que son especialmente aptas para la generación de hipótesis “en áreas de conocimientos no muy bien establecidos, donde es imposible contestar correctamente” (Ausubel: 533). Las prácticas profesionales tienen esa característica de intervenciones en situaciones donde lo imperante es la incertidumbre.

Las pruebas nacionales ponen énfasis en la medición de la denominada inteligencia escolástica, porque eso es lo que han requerido tradicionalmente las universidades. No obstante, la inteligencia escolástica no es el tipo de inteligencia que se requiere para alcanzar un desempeño más o menos exitoso en la vida y en las actividades científicas, profesionales y empresariales.

El criterio de validez, por lo tanto, debería ser que las evaluaciones midan lo que se pretende medir, esto es, el conocimiento y la reflexión y no la mera retención de información bruta.

Una opción difícil

Si entre un cuarenta y un cincuenta por ciento de los alumnos que rinden los exámenes nacionales de evaluación no logran reunir los puntos necesarios para ingresar en las universidades, es posible que no deban ingresar a los estudios superiores, desde el punto de vista del corto plazo de los recursos financieros. Pero la experiencia del examen debería servirles a los rechazados, por lo menos, para reanudar con una mejor

orientación sus exploraciones vitales y no para abandonarse a sí mismos luego de una primera frustración. Es importante que la discriminación que se realiza con los puntajes de las evaluaciones nacionales midan la capacidad de reflexión que se necesita para resolver problemas en los estudios superiores. Si la medición de esa capacidad es reemplazada por una acumulación de puntuaciones vinculadas con el material almacenado en la memoria de corto plazo y por una táctica orientada a acertar en las opciones múltiples (a la manera de los SAT), las evaluaciones nacionales solamente contribuirán a no solucionar la situación de los alumnos inertes, debido a la acumulación de información no operable y al refuerzo de la desmotivación que produce el fracaso escolar.

El compromiso de la elección es claro: las preguntas con opciones múltiples son más convenientes desde el punto de vista financiero-masivo, pero no miden lo que deberían medir. La técnica de las preguntas abiertas es más costosa para su compilación, pero mide lo que se debe medir.

Afortunadamente, unas palabras de prudencia nos llegan desde Howard Gardner, tal como fueron comentadas por el ex ministro Llach en una conversación personal en Harvard:

Gardner puso énfasis en la necesidad de enseñar la capacidad de comprensión. Enseñar a aprender. No a memorizar. No a llenarle la cabeza a los chicos con conocimientos para que los repitan mecánicamente. “Entonces surge la necesidad de revisar cómo se hacen las evaluaciones –opinó el ministro-. Muchas veces, las pruebas están diseñadas para medir cuántos datos saben repetir los chicos, más que comprobar si realmente han sido capacitados en los mecanismos de comprensión. [1]”

Notas:

Ausubel, David P.; Novak, Joseph D.; Hanesian, Helen: *Psicología Educativa: Un punto de vista cognitivo*. México: Trillas, 1997 (1968).

Cavanaugh, Sally Hixon: *Connecting Education and Practice* , en Curry , Lynn ; Wergin, John F.: *Educating Professionals* , Jossey-Bass Publishers, San Francisco, 1993:

Gardner, Howard: *Intelligence Reframed* . New York: Basic Books, 1999.

Llach, J.J.; Montoya S.; Roldán F.: *Educación para todos*. Córdoba, IERAL 1999.

Doerr Owen, David; Doerr, Marilyn: *None of the Above* . New York: Rowman & Littlefield Publishers, Inc., 1999.

Sternberg, Robert Ed.: *Thinking and Problem Solving*. San Diego (Cal.): Academic Press, 1994.

[1] "La Nación" del 17 de marzo de 2000