

Nivel de conocimientos y juicio reflexivo como indicadores de calidad de la formación superior en psicología. *

Francisco Leal Soto**, Roberto Storey Meza*** y Juan Carlos Díaz****.

Resumen

Se desarrolló y validó preliminarmente dos indicadores de evaluación de calidad en la formación psicológica, entendida como transformación o valor agregado al estudiante: el nivel de conocimientos específicos en Psicología, y el nivel de juicio reflexivo alcanzado por los estudiantes durante su formación.

Se utilizó un diseño cuasi-experimental transeccional, ya que se esperaba demostrar que los instrumentos discriminaban entre grupos de estudiantes de diferente nivel de avance en la carrera y de diferentes universidades, tomándose una muestra de estudiantes de primer, tercer y quinto año de la carrera de Psicología de tres universidades.

Esta investigación fue realizada en el marco del proyecto DIEXSA-3731-99 de la Universidad de Tarapacá, con aporte complementario de la Universidad José Santos Ossa (actualmente, parte de la sede Antofagasta de la Universidad del Mar). Un informe preliminar de este mismo trabajo fue presentado por los autores en las Jornadas Internacionales de Psicología Educacional "Psicología y Educación: encuentros y desencuentros", Universidad de La Serena, 18 al 20 de Octubre de 2001. Los autores agradecen la colaboración del Depto. de Psicología de la Universidad de La Serena.

Psicólogo, Depto. de Filosofía y Psicología, Universidad de Tarapacá.

Psicólogo, Depto. de Filosofía y Psicología, Universidad de Tarapacá.

Psicólogo, Escuela de Ciencias Sociales, Universidad del Mar, Antofagasta.

Los resultados permiten establecer que ambos indicadores son preliminarmente válidos, ya que distinguen entre niveles y universidades, y complementarios, ya que muestran un nivel moderado de correlación.

Abstract

Two indicators of quality in psychology education, whether as a transformation or as an added value to students, were developed and preliminarily validated: the level of specific knowledge in Psychology, and the level of reflective judgment level.

A semi-experimental, trans-sectional design was used, since it was expected to demonstrate that the instruments reveal differences among various level groups and among different universities. The participants were students from first, third and fifth grades in Psychology programs of three different universities.

The results allow us to establish that both indicators are preliminarily valid because they establish differences between levels and universities, and are also complementary as they show moderate level of correlation between them.

INTRODUCCIÓN

La calidad de la formación universitaria es un problema cada vez más urgente, dada la actual situación de la educación superior en Chile. La proliferación de universidades privadas y las decisiones acerca del financiamiento de las universidades públicas hacen que el desarrollo de sistemas o procedimientos de evaluación de la calidad sea urgente. Asimismo, la cada vez más cercana posibilidad de integración internacional a nivel profesional exige que en cada país se desarrollen procedimientos que aseguren el cumplimiento de ciertas normas o estándares básicos en la formación profesional.

Las primeras aproximaciones al tema han sido indirectas, evaluando la infraestructura, la implementación, la calidad del cuerpo docente y la eficiencia, bajo el supuesto de que existe una relación directa entre estos indicadores y la calidad de la formación impartida a los alumnos. Sin embargo, no se ha abordado directamente la evaluación del impacto de la formación universitaria en los alumnos, como se ha hecho en otros niveles del sistema educacional chileno con el SIMCE (Sistema de Medición de la Calidad de la Enseñanza),

que evalúa directa y masivamente el nivel de logro alcanzado por los alumnos en determinados cursos.

Más aún, una de las formas más recientes e integrales de entender la calidad en la educación superior se refiere a este concepto como transformación o valor agregado. Esta noción se basa en el cambio cualitativo y cuestiona la idea de calidad centrada en el producto, pues considera que la calidad radica, por un lado, en desarrollar las capacidades del estudiante y, además de ello, posibilitarle para influir en su propia transformación. En el primer caso, el «valor agregado» es una medida de calidad en términos del grado en que la experiencia educativa incrementa el conocimiento, las capacidades y las destrezas de los estudiantes. En el segundo, supone una implicación del estudiante en la toma de decisiones que afectan su transformación que, a la vez, proporciona la oportunidad de ampliar sus posibilidades para participar en la vida profesional (Rodríguez, 1997).

La calidad como experiencia de transformación es una nueva noción de calidad que encuentra sus raíces en el concepto de cambio cualitativo, refiriéndose a un cambio fundamental de forma, en que la educación resulta ser un proceso participativo, no un mero servicio para un cliente, sino un proceso continuo de transformación del participante, sea estudiante o investigador. En educación, la transformación cualitativa está compuesta por dos elementos: el mejoramiento del participante y su participación en la toma de decisiones, que consiste en darle poder para que influya sobre su propio proceso de transformación y se apropie del proceso de aprendizaje (Harvey, 1998).

En palabras de María José Lemaitre (2000), la calidad se juega en la capacidad de la educación superior para transformar a sus estudiantes, produciendo en ellos un cambio cualitativo, en que se modifican las características fundamentales de la relación del individuo con su medio. En este mismo sentido, Astin (en Rodríguez, 1997) afirma que la calidad de una institución universitaria radica en su capacidad para lograr el mayor desarrollo posible de sus miembros, tanto alumnos como profesores; es decir, en conseguir el máximo valor añadido. Entiende valor añadido como los diferenciales del grado de desarrollo intelectual, afectivo, personal y social de los alumnos entre el inicio y el final de sus estudios, así como el desarrollo de los profesores, tanto en su dimensión pedagógica como académica y de productividad científica.

En este trabajo nos propusimos avanzar hacia la evaluación de la calidad de la educación superior en Psicología, entendida como valor agregado,

desde dos aspectos. El primero es el logro de los conocimientos y habilidades específicas para el desempeño profesional en un área determinada, lo que podría relacionarse con los objetivos específicos del currículum. El segundo es el logro de aquellos objetivos generales, usualmente referidos a aspectos globales deseables para cualquier profesional universitario, como podrían ser la amplitud de criterio, la creatividad en la búsqueda de soluciones y la habilidad para adaptar el conocimiento a las situaciones específicas. Este segundo aspecto generalmente no se encuentra operacionalizado en los *currícula*, y en muchos ni siquiera se encuentra explicitado; forma parte del currículum implícito u oculto.

Resulta evidente que estos dos aspectos no son evaluables de la misma forma. El primero parece directamente evaluable porque se refiere a objetivos y contenidos, que además están generalmente explicitados y operacionalizados en los respectivos planes y programas de estudio; aunque esta evaluación puede ser problemática debido a que se requiere definir cuáles son los objetivos y contenidos relevantes de evaluar, sobre todo si la evaluación es interinstitucional -lo que implica eventualmente perspectivas teóricas, aproximaciones técnicas y áreas de aplicación distintas- se trata de una dificultad técnicamente soluble.

Distinto es el caso de las habilidades generales consideradas en el segundo aspecto. Como se ha dicho, muchas veces no están ni siquiera explicitadas en los planes de estudio, y cuando lo están, no se encuentran operacionalizadas. En tales circunstancias, ¿Cómo saber si se han logrado? ¿Cómo proponer un procedimiento de evaluación, si no se dispone de los criterios requeridos? En este caso, el problema no es simplemente técnico, ya que se requiere construir un marco que permita explicitar y operacionalizar los criterios que se encuentran implícitos en planes diversos que permiten, además, múltiples interpretaciones.

Las investigaciones sobre el desarrollo cognitivo en adultos aportan elementos relevantes para encarar la evaluación desde este segundo aspecto. Se ha mostrado que el desarrollo cognitivo continúa más allá de la adolescencia (King y Kitchener, 1994) y que se relaciona fuertemente con el acceso y aprovechamiento de la escolaridad superior (Schmidt, 1985; Welfel, 1982; Welfel y Davison, 1986). Kitchener y King (1981) han propuesto un modelo de desarrollo cognitivo durante la adolescencia y la adultez, que puede ser evaluado directamente (King y Kitchener, 1994) y cuya validez ha sido mostrada en estudiantes chilenas de Psicología (Leal, 1997). Según este modelo, el cono-

cimiento se da en tres niveles: cognitivo, metacognitivo y epistémico (Kitchener, 1983). El primer nivel ocurre cuando la persona se ocupa directamente en tareas que le permiten construir el conocimiento del mundo, tales como percibir, calcular, leer y memorizar. Tal tipo de tareas constituye el nivel cognitivo. Pero al involucrarse en esas tareas, la persona realiza simultáneamente un segundo tipo de procesos, que le permite darse cuenta de cómo realiza los primeros; por ejemplo, al intentar memorizar una lista, la persona puede darse cuenta de que si agrupa los elementos de alguna forma lógica, los recuerda más fácilmente, y entonces puede utilizar tal estrategia para facilitar su tarea o mejorar su ejecución. Estos procesos del segundo nivel constituyen la metacognición, que si bien ha sido definida en forma amplia como el «conocimiento y cognición acerca de los fenómenos cognitivos» (Flavell, 1979, p. 906, en Kitchener y King, 1994), se aplica en la práctica en forma restringida al monitoreo de los propios procesos cognitivos y de los factores que los afectan cuando se trabaja en una tarea u objetivo específico (Kitchener, 1983).

Atendiendo a este uso restringido del concepto de metacognición que se evidencia al revisar los estudios al respecto, lo que deja fuera del mismo la posibilidad de juicios valorativos, esta autora propone que ellos se dan en un tercer nivel, de «meta-metacognición», a través del cual se monitorea la naturaleza epistémica del problema y el valor de verdad de las soluciones alternativas. Este nivel se relaciona con el conocimiento acerca del conocimiento en un sentido propiamente epistémico, por lo que es llamado «cognición epistémica», e incluye el conocimiento acerca de los límites del conocer, de la certeza del conocimiento, de los criterios del conocimiento válido y de las estrategias para identificar las formas de solución requeridas por un problema y escoger entre ellas.

Kitchener (1983) señala que cuando las personas se enfrentan a un problema difusamente estructurado, deben preguntarse primero si el problema es soluble, y si lo es, en qué forma; asimismo, si hay estrategias disponibles para encontrar la solución; preguntas todas de carácter epistémico. Si la persona no se hace estas preguntas, tratará todos los problemas como si fueran bien estructurados, y no se involucrará en procesos verdaderamente reflexivos, en el sentido aquí expuesto. Así, el juicio reflexivo surge como posibilidad cuando la cognición epistémica adopta ciertas formas; es decir, cuando los supuestos de la persona acerca del conocimiento son de tal tipo que le permiten reconocer soluciones alternativas y escoger entre ellas. La posibilidad del juicio reflexivo que surge de una posición epistémica tal es relevante porque lleva a decisiones que son fruto de una opción razonada, de una elección deliberada

que no es posible cuando los supuestos epistémicos no permiten ver soluciones alternativas que pueden ser evaluadas como mejores o peores dado un marco de referencia, ni tampoco permiten ver la incerteza (Kitchener y King, 1981).

Las distintas formas en que se manifiesta la cognición epistémica siguen una secuencia evolutiva desde formas simples hasta otras más complejas, sofisticadas e integradas, de las que el juicio reflexivo sería el estadio de mayor complejidad; secuencia que Kitchener y King han descrito en lo que llaman el «Modelo de Juicio Reflexivo» (Kitchener y King, 1981, King y Kitchener, 1994). Según ellas, el modelo «describe la progresión evolutiva que ocurre entre la niñez y la etapa adulta en la forma en que las personas entienden el proceso de conocer y en las correspondientes formas en que justifican sus creencias acerca de los problemas difusamente estructurados. En otras palabras, el modelo describe el desarrollo de la cognición epistémica. A medida que el individuo se desarrolla, va siendo capaz de hacer mejores evaluaciones acerca de las afirmaciones cognoscitivas, y de explicar y defender mejor sus puntos de vista en asuntos controvertidos. La habilidad de hacer juicios reflexivos es el último resultado de esta progresión» (King y Kitchener, 1994, p. 13).

El modelo propuesto por Kitchener y King se organiza en siete estadios. Cada estadio sucesivo en este modelo implica una mejor y más efectiva forma de justificación, provista de supuestos más inclusivos e integrados para evaluar y defender puntos de vista; cada uno se caracteriza, además, por una diferente forma de enfrentar los problemas pobremente estructurados. Sin embargo, las diferencias entre los estadios no son igualmente importantes; puede diferenciarse tres grupos, en función del supuesto principal que subyace a la lógica epistémica en ellos.

Los tres primeros estadios, a pesar de sus diferencias, comparten la lógica básica de que el conocimiento se obtiene a través de la observación personal y directa, o a través del trabajo de los expertos; en otras palabras, la realidad existe objetivamente y puede ser efectivamente conocida, por lo que el conocimiento es correcto y verdadero o incorrecto y falso. Las personas que se ubican en estos estadios no reconocen los problemas difusamente estructurados, tratando todos los problemas de la misma manera: como problemas bien estructurados. Este tipo de pensamiento, correspondiente a los tres primeros estadios, es llamado por King y Kitchener (1994) «pensamiento pre-reflexivo». Los estadios 4 y 5 forman la siguiente unidad, caracterizada

por el reconocimiento de la incertidumbre, lo que permite reconocer los problemas pobremente estructurados, y, por tanto, que hay problemas en que efectivamente no se puede resolver con certeza; sin embargo, no hay forma de resolver la incerteza, por lo que la evaluación es, en última instancia, siempre idiosincrática; en otras palabras, se reconoce la validez de distintos puntos de vista, pero no se puede establecer juicios acerca de su valor de verdad más allá de su significación idiosincrática o en un contexto particular. Este razonamiento es caracterizado como «pensamiento cuasi-reflexivo». El pensamiento de los estadios 6 y 7 corresponde a lo que las autoras denominan propiamente «pensamiento reflexivo», y se caracteriza por un conjunto de suposiciones epistémicas basadas en que el conocimiento no es algo dado, sino que debe ser construido activamente por el sujeto, y debe ser entendido en el contexto de relaciones conceptuales en que es generado; por lo mismo, no hay una «verdad efectiva», pero se puede evaluar las distintas posiciones desde la perspectiva de la validez de las reglas sobre las cuales ha sido construido, por ejemplo, solidez del razonamiento, validez de los datos, coherencia, parsimonia de la explicación, generalidad, integralidad, etc. Esto implica un notable avance sobre el pensamiento pre-reflexivo, al proporcionar una salida al escepticismo y al «idiosincratismo» de dicha etapa.

El desarrollo cognitivo en la adolescencia y adultez se daría principalmente en el tercer nivel de este modelo, el nivel epistémico, dando origen a la posibilidad, en los adultos, del juicio reflexivo, última etapa de desarrollo en dicho nivel. La posibilidad de este desarrollo está fuertemente relacionada con la educación superior: al egresar de High School, la media corresponde al estadio 3 (nivel pre-reflexivo; King y Kitchener, 1994, p. 279); al terminar el pregrado, al estadio 4, y para los estudiantes de postgrado es el estadio 5 (ambos cuasi-reflexivos, Welfel y Davison, 1986; Montecinos, 1989). Este efecto de la escolaridad también ha sido mostrado en estudiantes chilenas (Leal, 1997). Por esta razón, podría ser un buen indicador directo del impacto de la educación superior.

Por otra parte, aunque la disciplina en que se recibe la educación superior no parece hacer diferencias en cuanto al nivel de desarrollo de la cognición epistémica (Welfel, 1982; Welfel y Davison, 1986), el desarrollo del juicio reflexivo no se presenta generalizadamente, sino más bien en relación a las áreas específicas de contenido a que el sujeto ha estado expuesto (Montecinos, 1989; Leal, 1997); es decir, su desarrollo resulta dependiente del contexto, siendo mayor en aquellas áreas en que el sujeto tiene información o códigos disponibles para contextualizar los problemas a los que debe aplicar el juicio

reflexivo. Esto hace al desarrollo del juicio reflexivo aún más interesante como posible indicador directo del efecto de la educación superior, ya que permitiría diferenciar el desarrollo a nivel global y en relación directa con la disciplina estudiada.

Metodología

En este trabajo se evaluó el nivel de conocimientos y el desarrollo de la cognición epistémica en estudiantes de tres niveles de la carrera de Psicología de tres universidades chilenas, a fin de determinar si estos indicadores establecían discriminaciones entre niveles y eventualmente entre universidades; en otras palabras, si resultaban indicadores válidos para evaluar diferencias entre cada uno de estos grupos. Para ello, se utilizó un diseño cuasi-experimental trans-seccional, ya que se esperaba demostrar que los instrumentos discriminaban entre grupos de estudiantes de diferente nivel de avance en la carrera. La calidad cuasi-experimental del diseño está determinada por las limitaciones del muestreo y por las características de la variable independiente -años de estudios universitarios- que no permite asignación aleatoria, actuando como una variable ex-post-facto.

La muestra fue de participación voluntaria, por lo que su constitución no fue manejada por los investigadores. Se diseñó para disponer de 9 grupos de alumnos de la carrera de Psicología, estratificados por universidad (3 universidades, dos estatales derivadas y una privada) y por año de ingreso (para incluir alumnos de 1º, 3º y 5º año). La constitución final de la muestra, luego de descartar los casos que presentaron información incompleta, fue la contenida en la Tabla N° 1.

Para evaluar el nivel de conocimientos adquiridos por los alumnos, se desarrolló un instrumento específico, consistente en una prueba de ítemes de respuesta cerrada, del tipo de selección múltiple. Para ello, en primera instancia se estableció las áreas y contenidos fundamentales de estudios para los programas de Psicología, basándose en las recomendaciones del Colegio de Psicólogos de Chile y la Red Nacional de Unidades Académicas de Psicología (Colegio de Psicólogos de Chile, 1996), en el trabajo de Amalio Blanco (Blanco, 1995) y en los programas de estudio de dos de las universidades involucradas. El resultado, contenidos mínimos comunes para ambas universidades, fue sometido a consulta con académicos de las mismas, y modificado de acuerdo a las sugerencias recibidas. Sobre esa base, se solicitó a los académicos de ambas universidades elaborar ítemes que

midieran el dominio de esos contenidos, disponiéndose de una base de aproximadamente 120 ítemes que cubrían dichas materias. De esos ítemes se seleccionó, con criterios de claridad, representatividad de los contenidos e importancia relativa, 75, que constituyeron el instrumento que finalmente se aplicó a los sujetos.

Luego de la aplicación, los resultados se sometieron a análisis de ítemes y confiabilidad, eliminándose aquellos que mostraron escasa dispersión (aquellos muy fáciles o muy difíciles) o cuya correlación con el puntaje total era de tal naturaleza que disminuía la confiabilidad del instrumento. Se obtuvo finalmente una escala con 56 ítemes y una confiabilidad por consistencia interna (alpha de Cronbach) de 0.87, cuya validez de contenido estaba asegurada por el procedimiento de construcción.

Para evaluar el desarrollo de la cognición epistémica se utilizó una versión ligeramente modificada del Test Escrito de Juicio Reflexivo desarrollado por Mosquera y Murillo (2000), a través de una actividad de graduación desarrollada como parte de este proyecto. El Test Escrito consiste en dos temas controversiales -la legalización de la marihuana en Chile y la responsabilidad gubernamental o civil en el manejo de la delincuencia- planteados brevemente, después de cada uno de los cuales se pide a los sujetos responder por escrito, en un pequeño espacio asignado, a 5 preguntas orientadas a elicitare respuestas que pudieran ser puntuadas desde el Modelo de Juicio Reflexivo de King y Kitchener (1994).

El Test Escrito fue validado en forma piloto por Mosquera y Murillo en un grupo de 37 estudiantes universitarios de primer, tercer y quinto nivel de la Universidad de Tarapacá (Arica, Chile), de cuatro carreras de distintos ámbitos (Obstetricia, Derecho, Pedagogía en Inglés e Ingeniería), a quienes se aplicó el Test Escrito y la Entrevista de Juicio Reflexivo en la versión adaptada por Leal (1997). Ambos instrumentos fueron puntuados separadamente por dos evaluadores, alcanzándose una confiabilidad interjueces de 0,79 (Coeficiente de concordancia W de Kendall) en el caso del Test Escrito. La validez se obtuvo correlacionando el nivel promedio asignado a cada sujeto por ambos evaluadores en los dos instrumentos, a través del coeficiente de correlación Rho de Spearman, dado que el nivel de medición es caracterizado más apropiadamente como ordinal. El resultado fue una correlación de 0,75, lo que indica un razonable grado de validez concurrente.

Se registró datos de otras variables con fines de descripción de la muestra y análisis complementarios para mejor interpretar los resultados obtenidos en las variables principales. Las variables adicionales registradas fueron: universidad, edad, sexo, número de asignaturas reprobadas, grupo de rendimiento en relación a su promoción en que se autopercibe el alumno (tercio de menor, medio o mayor rendimiento), dependencia del establecimiento en que cursó la enseñanza media (municipal, particular subvencionada o particular pagada), nivel educacional de los padres (el mayor de ambos), y cinco variables referidas a las respuestas al Test Escrito de Juicio Reflexivo. Estas variables corresponden a apreciaciones hechas por los correctores acerca de distintos aspectos de las respuestas, que se agregaron como consecuencia de la variabilidad que se observó en dichos aspectos, por lo que se pensó que podrían influir en los resultados. Su evaluación, por ser una mera apreciación concordada en base a criterios mínimos, es muy precaria. Estas variables y las categorías en que fueron valoradas son las siguientes:

Caligrafía: ilegible, legible con dificultad y fácilmente legible

Ortografía: presencia de errores graves que alteran el sentido de la oración, presencia de errores menores que no alteran el sentido de la oración y sin errores.

Construcción Gramatical: presencia de errores graves como inconsistencias y desorganización, presencia de errores menores y buena construcción gramatical.

Redacción: confusa, incomprensible o contradictoria; comprensible con esfuerzo; fácilmente comprensible; clara, comprensible y precisa.

Contenido de la Respuesta: orientada al contenido u orientada a los aspectos epistémicos consultados. Esta variable fue incorporada debido a que las respuestas orientadas a los aspectos epistémicos resultan más fácil y claramente puntuables.

Los instrumentos fueron aplicados a los sujetos y los datos codificados e introducidos a una base de datos y procesados estadísticamente con el programa SPSS (r), versión 8.0 (SPSS Inc., 1998). De acuerdo a los objetivos e hipótesis, se realizó análisis descriptivos, correlacionales y pruebas de significación de diferencias.

Resultados

El nivel de conocimientos fue significativamente diferente para cada año de ingreso, como se ve en las Tablas N°2 y N°3, y también mostró diferencias entre dos de las tres universidades, como se ve en las Tablas N°4 y N°5, aunque no se encontró efecto de interacción entre ambas variables.

En el nivel de juicio reflexivo resultaron significativos el efecto del año de ingreso y la interacción entre año de ingreso y universidad, como se muestra en el Anova presentado en la Tabla N°6; el efecto principal de la universidad no resultó significativo por sí mismo. El efecto de la interacción se aprecia claramente en el Gráfico N°1, en que se puede ver que la Universidad B, siendo la que presenta un menor nivel promedio de juicio reflexivo al inicio de la carrera, muestra un aumento muy significativo en el nivel promedio entre el 3° y el 5° año, a diferencia de las otras dos universidades, que muestran un menor aumento global, particularmente entre el 3° y el 5° año. Las diferencias son significativas entre 1° versus 3° y 5° en la Universidad A, entre 1° y 3° versus 5° en la Universidad B, y sólo entre 1° y 5° en la Universidad C.

Discusión

Estos resultados preliminares son satisfactorios, ya que, a pesar de la limitación impuesta por los instrumentos, que son aún muy groseros, permiten establecer distinciones claras por año de ingreso a partir de ambas variables. Estos resultados son más claros en el caso del nivel de conocimiento que en el de juicio reflexivo. El primero distingue en forma neta por año de ingreso, y en menor medida, también entre universidades, por lo que parece ser un buen indicador de diferencias entre grupos por universidad y año de ingreso.

Los resultados para el nivel de juicio reflexivo son igualmente alentadores, aunque es necesario considerarlos con más cautela. Al interactuar ambas variables -año de ingreso y universidad- aunque la tendencia es la misma en las tres universidades, es decir, a aumentar el nivel de juicio reflexivo a medida que aumenta la permanencia en la universidad, la magnitud de este aumento es diferente en cada una en los distintos niveles.

Si las diferencias encontradas reflejaran efectivamente diferencias en los procesos, con el resultado de agregar cuotas diferentes de valor a cada

grupo, estaríamos en condiciones de utilizar estos indicadores como medida razonable de diferencias en calidad educativa, tanto para establecer comparaciones intra como entre instituciones. Sin embargo, el diseño de este estudio no permite establecer esta conclusión, ya que al ser trans-seccional, no controla el efecto de cohorte, por lo que no sabemos si las diferencias se deben a diferencias iniciales entre los grupos de sujetos de un mismo año de ingreso, o a los procesos ocurridos durante su formación. De hecho, en las variables que se recogieron adicionalmente, se observó diferencias significativas entre universidades en las variables Asignaturas Reprobadas en los Años de Ingreso 1996 ($F=17.59$, $gl=2$, $\alpha=0.000$) y 1998 ($F=6.318$, $gl=2$, $\alpha=0.004$); Estudios de los Padres en el Año de Ingreso 1996 ($F=6.570$, $gl=2$, $\alpha=0.003$); Ortografía en el Año de Ingreso 2000 ($F=4.757$, $gl=2$, $\alpha=0.012$); Construcción Gramatical en los tres Años de Ingreso ($F=5.597$, $gl=2$, $\alpha=0.007$; $F=15.271$, $gl=2$, $\alpha=0.000$; y $F=17.896$, $gl=2$, $\alpha=0.000$, respectivamente); y Orientación de la Respuesta también en los tres Años de Ingreso ($F=25.445$, $gl=2$, $\alpha=0.000$; $F=7.628$, $gl=2$, $\alpha=0.001$; y $F=7.966$, $gl=2$, $\alpha=0.001$, respectivamente), por lo que queda claro que los grupos difieren en algunas características importantes que hacen difícil interpretar las diferencias en las variables dependientes. La única forma de establecer con claridad este punto es con un estudio de cohorte, para lo cual se está realizando una segunda evaluación a las mismas cohortes, de la cual aún no se tiene datos suficientes.

Otro elemento central en esta discusión lo constituyen la calidad y viabilidad de los instrumentos desarrollados para la evaluación del impacto de la educación superior en Psicología. Respecto de la prueba de conocimientos, resulta evidente la dificultad para elaborar un instrumento que evalúe válidamente la amplia variedad de contenidos incluidos en las distintas universidades. Si bien se estima que se logró construir un instrumento válido para las universidades participantes, dicha validez puede ser diferente para cada caso; este problema sólo puede ser superado en la medida en que efectivamente la validez del contenido del instrumento es cotejada con el plan y los expertos de cada universidad.

En el caso del instrumento para evaluar juicio reflexivo, la principal dificultad radica en la validez de la forma de evaluación, ya que se ha tratado de realizar por escrito una evaluación que originalmente se realiza a través de una entrevista semiestructurada, forma que permite una exploración en profundidad con cada sujeto hasta obtener respuestas clara y válidamente puntuables según el modelo. Al trasladar la evaluación al formato escrito, a través de un cuestionario de respuesta abierta, la posibilidad de interacción

entre evaluador y evaluado se pierde, y la respuesta no puede ser explorada; ello obliga a puntuar con respuestas que son, en general, mucho más pobres y menos claras que las obtenidas en la entrevista. Por ello, los indicadores psicométricos del instrumento desarrollado, siendo aceptables, no alcanzan niveles comparables a los obtenidos en la entrevista, lo que hace mella tanto a la validez como a la confiabilidad. Esta es una dificultad no menor, ya que el formato original resulta muy oneroso para ser utilizado masivamente; y los resultados obtenidos con el instrumento modificado resultan menos claros. En esta disyuntiva, se estima que lo alcanzado con el formato de respuesta abierta proporciona una alternativa transitoria con la cual seguir trabajando, mientras se desarrolla un formato más preciso y válido que sea al mismo tiempo económico, o bien mientras se busca un modelo alternativo cuya forma de evaluación sea menos compleja.

Mientras tanto, ambos instrumentos parecen proporcionar resultados suficientemente promisorios en términos de posibilitar la evaluación de cambios efectivos en los alumnos, de incorporación de valor agregado en los mismos, como una dimensión de calidad de educación superior.

Tabla N° 1:

Constitución final de la muestra por Universidad, Año de Ingreso y Sexo (Frecuencias)

Universidad			SEXO		Total
			Masculino	Femenino	
Universidad A	Año de ingreso	1996	7	13	20
		1998	8	17	25
		2000	11	13	24
		Total	26	43	69
Universidad B	Año de ingreso	1996	2	9	11
		1998	6	13	19
		2000	6	16	22
		Total	14	38	52
Universidad C	Año de ingreso	1996	2	10	12
		1998	1	8	9
		2000	2	10	12
		Total	5	28	33

Tabla N° 2:

Análisis de varianza para la Prueba de Conocimientos según Año de Ingreso
(Preguntas correctas prueba final)

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	5217.442	2	2608.721	62.464	.000
Within Groups	6389.808	153	41.763		
Total	11607.250	155			

Tabla N° 3:

Comparaciones Post-Hoc (Scheffé) para la Prueba de Conocimientos según
Año de Ingreso (Preguntas correctas prueba final)

(I) Año de ingreso	(J) Año de ingreso	Mean Difference(I-J)	Std. Error	Sig.	95% Interval Confidence	
					Lower Bound	Upper Bound
1996	1998	6.6736*	1.3208	.000	3.4085	9.9386
	2000	14.3528*	1.2958	.000	11.1497	17.5559
1998	1996	-6.6736*	1.3208	.000	-9.9386	-3.4085
	2000	7.6792*	1.2171	.000	4.6707	10.6877
2000	1996	-14.3528*	1.2958	.000	-17.5559	-11.1497
	1998	-7.6792*	1.2171	.000	-10.6877	-4.6707

* The mean difference is significant at the .05 level.

Tabla N° 4:

Análisis de varianza para la Prueba de Conocimientos según Universidad
(Preguntas correctas prueba final)

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	668.437	2	334.219	4.675	.011
Within Groups	10938.813	153	71.496		
Total	11607.250	155			

Tabla N° 5:

Comparaciones Post-Hoc (Scheffé) para la Prueba de Conocimientos según Universidad (Preguntas correctas prueba final)

		Mean	Std.Error	Sig.	95%	Interval
		Difference(I-J)			Confidence	
(I) Universidad	(J) Universidad				Lower Bound	Upper Bound
Universidad A	Universidad B	4.6736*	1.5433	.012	.85868	.4886
	Universidad C	1.2744	1.7814	.775	-3.1291	5.6780
Universidad B	Universidad A	-4.6736*	1.5433	.012	-8.4886	-.8586
	Universidad C	-3.3992	1.8819	.199	-8.0510	1.2526
Universidad C	Universidad A	-1.2744	1.7814	.775	-5.6780	3.1291
	Universidad B	3.3992	1.8819	.199	-1.2526	8.0510

* The mean difference is significant at the .05 level.

Tabla N° 6:

Análisis de varianza para Nivel de Juicio Reflexivo según Universidad y Año de Ingreso (Nivel de Juicio Reflexivo)

Source	Type III Sum of Square	df	Mean Square	F	Sig.
Corrected Model	34.750	8	4.344	6.403	.000
Intercept	1656.620	1	1656.620	2441.810	.000
UNIV	1.0722	.536	.790	.456	
INGRESO	25.438	2	12.719	18.747	.000
UNIV * INGRESO	8.846	4	2.211	3.260	.014
Error	99.731	147	.678		
Total	1983.000	156			
Corrected Total	134.481	155			

a R Squared = .258 (Adjusted R Squared = .218)

Referencias

Harvey, L. (1998) *La Educación en el Siglo XXI*. Revista Calidad en la Educación. Consejo Superior de Educación (CSE), Santiago, diciembre 1998.

- King, P. & Kitchener, K. (1994) *Developing Reflective Judgment*. San Francisco: Jossey-Bass Publishers.
- Kitchener, K. & King, P. (1981) *Reflective Judgment: Concepts of Justification and Their Relationship to Age and Education*. Journal of Applied Developmental Psychology: 2, 89-116.
- Kitchener, K. (1983) *Cognition, Metacognition, and Epistemic Cognition. A Three-Level Model of Cognitive Processing*. Human Development: 26, 222-232.
- Leal, F. (1997) *El Juicio Reflexivo en estudiantes de la Universidad de Tarapacá: un estudio trans-seccional*. Arica: Universidad de Tarapacá, Tesis inédita para optar al grado de Magíster en Educación.
- Lemaitre, M.J. (2000) *La Calidad en la Formación de Pregrado: Autoevaluación y Acreditación*. Revista Estudios Sociales. Corporación de Promoción Universitaria (CPU), Santiago, N° 103
- Montecinos, C. (1989) *A cross-sectional Study of the Development of Reflective Judgment Among Medical Students*. Carbondale: Southern Illinois University, Department of Educational Psychology. Tesis doctoral inédita.
- Personnel: September, 388-394.
- Rodríguez, Sebastián. (1997) *Orientación Universitaria y Evaluación de la Calidad*. En: Calidad en la Universidad. Apodaca, P., y Lobato, C. Eds. Barcelona: Laertes.
- Schmidt, J. (1985) *Older and Wiser? A Longitudinal Study of the Impact of College on Intellectual Development*. Journal of College Student Personnel: September, 388-394.
- Welfel, E. & Davison, M. (1986) *The Development of Reflective Judgment During the College Years: A 4-Year Longitudinal Study*. Journal of College Student Personnel: May, 209-216.
- Welfel, E. (1982) *How Students Make Judgments: Do Educational Level and Academic Major Make a Difference?* Journal of College Student Personnel: November, 490-496.