

EVALUACIÓN DE IMPACTO DE LAS
PRÁCTICAS PROFESIONALES Y
TRABAJOS DE GRADO EN UNA
INSTITUCIÓN DE EDUCACIÓN
SUPERIOR

Carlos A. Pardo Adames*, Karen Cabarcas y
Marcela Velasco.

Universidad Católica de Colombia, Bogotá, Colombia.

Introducción

A pesar de que la evaluación de impacto ha sido empleada en los últimos años debido a que proporciona información importante para valorar la efectividad de programas o procesos –lo que la ha hecho necesaria en los contextos educativos en general como soporte y aporte del diseño de políticas sectoriales a nivel institucional, local o nacional–, aún es una metodología nueva para implementar en la Universidad Católica de Colombia (o, por lo menos, poco frecuente), y no ha sido utilizada en los procesos de desarrollo o de implementación de procesos de formación de estudiantes en la Facultad de Psicología.

En términos generales, la evaluación de impacto se utiliza para conocer y valorar los efectos causales de proyectos o programas de intervención (IEG, 2012). Por ejemplo, el Independent Evaluation Group [IEG] (2012), del Banco Mundial, dice que la evaluación de impacto evalúa los efectos causales (impactos) atribuibles a las intervenciones específicas en donde los resultados de interés de los participantes de un programa se comparan con una situación contraria a los hechos (*counterfactual*²). Asimismo, Kandkher, Koolwal y Samal (2010) plantean que la evaluación de impacto estudia si los cambios se deben al programa de intervención o a otros factores; en el mismo sentido, Aldercotte (2018) indica que con esta estrategia se evalúa el impacto de políticas o procesos existentes o propuestos en relación con sus consecuencias; y Misra (s. f.) define el impacto social como las consecuencias que tiene en las poblaciones humanas cualquier intervención pública o privada que altere las formas en

*capardo@ucatolica.edu.co

2 El término *counterfactual* se ha traducido como “contrario a los hechos” debido a que se refiere a un grupo de personas que no ha pasado por el programa y que, por lo tanto, puede dimensionar –por contraste– el efecto del programa.

que una persona vive, trabaja o juega –entre otras– y, en general, en sus relaciones con los demás. Desde esta perspectiva, una evaluación de impacto se interesaría más por la pregunta de si un proceso, política o programa es efectivo. Adicionalmente, existen estudios (Legovini, Di Maro & Piza, 2015) que plantean que la evaluación de impacto también podría responder la pregunta sobre cómo hacer efectivo un proceso, una política o un programa.

Modelo de evaluación de impacto

En el marco del presente trabajo, en concreto se buscó evaluar el impacto de los programas de (a) Prácticas de formación y profesionales (pregrado y postgrado); (b) Seminario de Trabajo de Grado (pregrado y postgrado); y (c) Egresados; de la Facultad de Psicología de la Universidad Católica de Colombia.

Para la evaluación del impacto en cada caso, y teniendo en cuenta que se hará una evaluación de monitoreo y evaluación de impacto propiamente dicha, los pasos a seguir en la investigación son:

- a. *Indicadores*: (a) *Indicadores intermedios* para cada uno de los programas a evaluar, que miden tanto los insumos como los productos; y (b) *Indicadores finales*, que miden los resultados logrados con el programa. La Figura 7, tomada y adaptada de IEG (2012), muestra el marco de la evaluación de impacto en general, incluyendo el monitoreo.

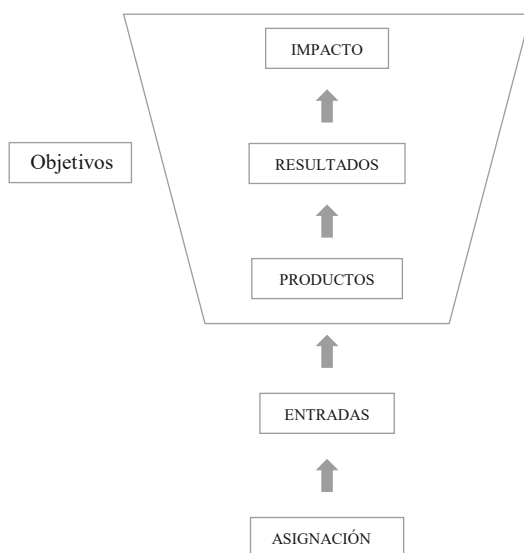


Figura 7. Esquema evaluación de impacto IEG (2012), p. 81.

- b. *Evaluación operacional de los programas*: consiste en el análisis de la implementación de los programas, con el objetivo de comprender si se desarrollan o ejecutan como se ha planeado. Es una evaluación retrospectiva y se hace con base en los objetivos planteados inicialmente (IEG, 2012). Durante varios años se planteó como alternativa una evaluación de impacto de tipo cuantitativa y de la forma *ex post*.
- c. *Selección del grupo de comparación*: en el caso de los programas de la facultad es muy difícil la selección de un grupo que reciba los programas y de uno contrario a los hechos, debido a que no es posible contar con un grupo de estudiantes que no hayan cursado los programas. En algunos casos es posible una evaluación de impacto “reflexiva” (IEG, 2012), es decir, tomando medidas antes y después, para lo que se requiere incluir estudiantes que aún no han iniciado los programas que se van a evaluar. Por ello, algunos de los estudiantes incluidos en este trabajo se encontraban en la mitad de implementación del programa y otros al final.
- d. *Selección de la metodología*: dadas las condiciones, se hizo una comparación de los resultados entre los grupos de estudiantes seleccionados.

De acuerdo con Gertler, Martínez, Premand, Rawlings y Vermeersch (2007), la evaluación de impacto responde esencialmente a preguntas de investigación que pretenden determinar la relación causa-efecto. En este sentido, la pregunta planteada para el presente trabajo fue: ¿cuál es el impacto (efecto) de las prácticas y el seminario de trabajo de grado en las habilidades, competencias y conductas de los estudiantes que cursan las asignaturas seleccionadas? De acuerdo con estos autores, la evaluación realizada fue retrospectiva, porque algunos estudiantes ya estaban asignados a una práctica y a un trabajo de grado, algunos llevaban más de un semestre y otros no habían empezado ni trabajo de grado ni prácticas (eran estudiantes de séptimo semestre que actuaron como el grupo contrario a los hechos o, por lo menos, uno de estos). Adicionalmente, se hizo la evaluación pre y post de los mismos estudiantes (de octavo y noveno semestre), en donde la evaluación pre identificaba a un grupo contrario a los hechos.

Teniendo esto en cuenta, se utilizaron medidas de comparación de medias entre los estudiantes de séptimo semestre y los de octavo y noveno semestre, considerando que la intervención (trabajo de grado y prácticas) no había ocurrido para los del primer grupo, lo que explicaría un desempeño inferior. Adicionalmente, se establecieron los pesos de regresión para cada alternativa de cada intervención (trabajo de

grado y prácticas) con el fin de comparar la efectividad de cada una de ellas y su significatividad en el desempeño de los estudiantes.

Diseño del instrumento

Con el fin de recolectar la información necesaria para la evaluación de impacto se diseñó un instrumento orientado a la autoevaluación del estudiante en relación con el desarrollo de competencias, habilidades y desempeños como efecto de su paso por los programas de trabajo de grado y de prácticas profesionales. El diseño del instrumento se realizó siguiendo la metodología centrada en evidencias (ECD, por sus siglas en inglés) de Mislevy, Almond y Lukas (2003), que representa una aproximación para construir evaluaciones en términos de evidencias. En este sentido, se tuvieron en cuenta los modelos planteados por los autores, los cuales se pueden observar en la Figura 8.

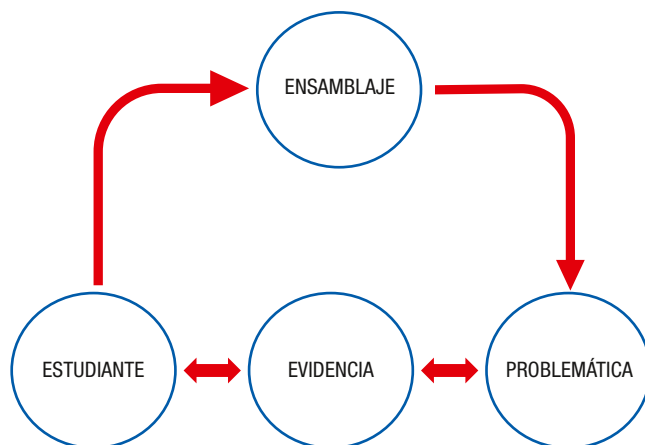


Figura 8. Elementos o modelos del marco de evaluación conceptual³.

Inicialmente, y siguiendo la metodología, se realizaron cuadros resumen de competencias, habilidades y conocimientos para cada una de las modalidades de trabajo de grado y de prácticas profesionales. A continuación se describe la información generada.

Indicadores para el instrumento

Con base en la información de las competencias, habilidades y conocimientos de cada alternativa de trabajo de grado y cada práctica profesional, así como de las competencias generales y específicas para el trabajo de grado y las prácticas profesionales, se elaboraron los indicadores a tener en cuenta en la evaluación de impacto y los

³ Tomado de la conferencia de Carlos Pardo en el contexto de “Jueves de Extensión”, de la Facultad de Psicología de la Universidad Católica de Colombia, en agosto de 2012.

ítems que componen la prueba utilizada. Los indicadores se agruparon en seis dimensiones que posibilitan el análisis de resultados:

- a. Formulación de acciones profesionales relacionadas con la investigación o el ejercicio de la psicología.
- b. Actividades orientadas a la selección de estrategias relacionadas con la investigación o el ejercicio de la psicología.
- c. Identificación de problemáticas relacionadas con la investigación o el ejercicio de la psicología.
- d. Comunicación de resultados relacionados con la investigación o el ejercicio de la psicología.
- e. Aplicación de procesos relacionados con la investigación o el ejercicio de la psicología.
- f. Grado de desarrollo de habilidades y competencias relacionadas con la investigación o el ejercicio de la psicología.

Los resultados permiten establecer, en detalle, el impacto del proceso de formación de los estudiantes y de las asignaturas de trabajo de grado y prácticas. Finalmente, de acuerdo con las categorías de instrumentos psicométricos propuesta por Aiken (1996), este instrumento corresponde a la clase de escalas de valoración (*rating scales*), ya que se quiere que cada estudiante valore –en una escala de 1 a 10 puntos, siendo el 10 el de mayor valor– una serie de conductas que representan las habilidades y competencias que debe desarrollar en las asignaturas de trabajo de grado y práctica.

Resultados

Los resultados se pueden clasificar en dos categorías: la primera, relacionada con la calidad técnica del instrumento utilizado, y considerando sus características; y, la segunda, que hace referencia a los resultados de impacto encontrados.

Calidad técnica del instrumento utilizado

Como se mencionó previamente, el instrumento consta de seis dimensiones y un total de 30 ítems en donde los estudiantes expresan el grado en el que consideran (auto-percepción) que tienen o han desarrollado una habilidad o competencia, o en el que pueden realizar alguna actividad puntual relacionada con las competencias objeto de evaluación.

Población

La población utilizada para el análisis técnico del cuestionario fue de 192 estudiantes de séptimo a noveno semestre (antes y durante el curso de las asignaturas correspondientes a trabajo de grado y prácticas profesionales). El cuestionario se aplicó a 204 estudiantes, pero se eliminaron datos de 12 de ellos (véase Tabla 8), debido a un proceso de valoración poco confiable (dar la misma respuesta a todas las preguntas o ítems), por errores en la digitación de los datos, o por ausencia de respuestas. En la Tabla 8 se encuentra el número total de estudiantes según semestre y jornada.

Tabla 8. Número de estudiantes según semestre y jornada

Semestre	Jornada	N
7	1	49
8	0	1
8	1	7
8	2	53
9	1	4
9	2	78
Total		192

Tabla 9. Distribución de estudiantes según trabajo de grado y práctica profesional

Alternativa de trabajo de grado	Semestre	N	Sin categorías	Clínica	Organizacional	Educativa	Comunitaria	Jurídica
Sin dato	7	49	49					
Sin dato	8	6	1	2	2		1	
Sin dato	9	2		1			1	
Curso de Especial Interés	8	20		11	7		2	
Curso de Especial Interés	9	34		11	10	3	3	7
Investigaciones	8	32		21	5		6	
Investigaciones	9	33		13	9	4	3	4
Servicio Social Comunitario	8	1					1	
Servicio Social Comunitario	9	9		3	1	3	1	1
Movilidad	8	2		1	1			
Movilidad	9	4		2	1			1
Total		192	50	65	36	10	18	13

En la Tabla 9 se reportan resultados con respecto a la distribución de estudiantes en cada alternativa de trabajo de grado y práctica profesional, la columna “sin categorías” refiere a los estudiantes de séptimo semestre que no han iniciado ninguno de los dos procesos, pero que son la base de datos de estudio y comparación a futuro (grupo contrario a los hechos). En las filas, el reporte “sin dato” corresponde a aquellos estudiantes que no marcaron alternativa o práctica, y que fueron tratados como dato perdido. De igual forma, en esta tabla se reporta un total de 192 estudiantes, de los cuales la mayor parte hace prácticas en las áreas clínica y organizacional; y la mayoría hace trabajo de grado en investigaciones y curso de especial interés (CEI). Cabe resaltar que el dato de estudiantes en la alternativa de movilidad es bajo, debido a que en el año 2016 se registraron pocos estudiantes –por ello, hay que tener en cuenta esta insuficiencia en los datos y, por tanto, en los resultados–.

Modelo utilizado

Para evaluar la calidad técnica del instrumento diseñado se utilizó el Modelo de Rasch de la Teoría de Respuesta al Ítem. Algunos autores (Chong, 2010) plantean que el Modelo de Rasch corresponde a un tipo de formulación diferente al de la Teoría de Respuesta al Ítem (TRI). No obstante, en términos generales, la TRI es una aproximación conceptual y técnica al uso y análisis de datos provenientes de pruebas o instrumentos de recolección de información con ciertas características. En esencia, la TRI corresponde a una aproximación diferente en muchos aspectos respecto al tipo de aproximación de la denominada Teoría Clásica de los Test (TCT), ya que autores como Jones y Thissen (2007) plantean que la TCT realiza inferencias del constructo evaluado a partir del ejercicio de la persona en toda la prueba (la puntuación total), mientras que la TRI realiza inferencias a partir del desempeño del evaluado en los diferentes ítems que aborda.

En particular, el análisis de Rasch es un método para obtener medidas objetivas, fundamentales y aditivas a partir de observaciones estocásticas de categorías de respuesta ordenadas; estas medidas se encuentran cualificadas por estadísticas de ajuste, lo que hace de este un modelo prescriptivo (Linacre, 2013).

De acuerdo con Linacre (2013), el software Winsteps es un programa de computador que realiza numerosas aplicaciones del modelo de Rasch, dentro de las cuales se encuentra el de la escala de valores, o valoración de Andrich, que corresponde a la clasificación del instrumento utilizado y descrito con anterioridad. Específicamente, el modelo de escala de valoración que maneja el software Winsteps (Linacre, 2013) se expresa como:

$$\log(P_{nij} / P_{ni(j-1)}) = B_n - D_i - F_j \quad (2)$$

Donde:

P_{nij} = La probabilidad de respuesta de una persona a un ítem particular en una de las categorías de valoración de la escala (que en el caso del presente instrumento se encuentra en una escala de 10 puntos).

B_n = La habilidad de la persona, entendida en este caso como la percepción de desarrollo de las habilidades o competencias o acciones evaluadas.

D_i = La complejidad relativa de los ítems planteados en términos de la facilidad para lograr lo planteado.

F_j = La medida calibrada de los distintos valores de cada ítem. No hay restricción en el número de categorías.

Indicadores de ítems

Este análisis técnico corresponde a un proceso cuantitativo y cualitativo mediante el cual se establece la calidad técnica de los ítems y del instrumento en relación con los propósitos para los cuales fueron elaborados y considerando los criterios matemáticos y estadísticos establecidos por el modelo, la experiencia y los estándares de diseño de instrumentos. Dado que se requiere procesar la información (respuestas a los ítems) para obtener los diferentes indicadores, fue necesario depurar la base de datos inicial en consideración a diferentes aspectos que podrían introducir ruido en las interpretaciones y valoraciones del instrumento.

Los indicadores propuestos en el procesamiento de respuesta son los más ampliamente utilizados en distintos estudios de evaluación educativa y se corresponden con los planteamientos de AERA, APA y NCME (1999), Pardo, Rocha y Barrera (2005), y Embretson y Hershberger (1999), y estos son:

- a. *Ajuste próximo y lejano (Infit y Outfit)*: indica la correspondencia entre un grupo de datos y el modelo estadístico utilizado para representarlos. El ajuste próximo (*Infit*) se refiere a la relación entre los datos que se encuentran cerca del valor de dificultad del ítem y el valor de dificultad; mientras que el ajuste lejano (*Outfit*) se refiere a la relación de los datos que se encuentran lejos de dicho valor de dificultad. *Interpretación*: los valores posibles se encuentran entre cero e infinito positivo; el valor que determina el ajuste perfecto entre los datos y el modelo es 1; los valores muy inferiores a 1 indican dependencia de los datos (paradoja de atenuación); los valores superiores a 1 indican ruido en la información; y valores de ajuste inferiores a 1.5. indican la posición de la curva del ítem a lo largo de la escala de habilidad, pues entre más difícil es un ítem, su curva estará localizada más a la derecha en la escala de habilidad. Este aspecto no tiene un criterio de aceptación o rechazo (Pardo & Rocha, 2010).
- b. *Correlación producto momento punto-medida (rpbis)*: indica la relación entre la respuesta correcta a un ítem de una prueba y el valor de habilidad obtenido en dicha prueba. Se calcula para el ítem en su totalidad (considerando el aporte de

- las categorías de respuesta según el modelo seleccionado). *Criterio de análisis*: es deseable que la distribución de valores de correlaciones de los ítems de la prueba incluya valores superiores 0.25 (Pardo & Rocha, 2010).
- c. *Indicadores de prueba*: para el análisis de la prueba en su totalidad se seleccionaron dos indicadores: la confiabilidad y el resultado del análisis de componentes principales que trae el software para identificar si las dimensiones en realidad son dimensiones (Pardo & Rocha, 2010).
 - d. *Confiabilidad*: este concepto se puede entender de muchas formas (Pardo & Rocha, 2004): la primera, como el grado en el cual un instrumento de medición mide algo en forma consistente –es decir, que un instrumento confiable está relativamente libre de errores de medición–; la segunda, como el grado en el cual los puntajes –o cualquier otra forma de resultado– obtenidos en una forma específica de prueba, administrada en un momento particular y posiblemente puntuada por alguna persona en particular, puede ser generalizada a puntajes obtenidos en otras formas de prueba, administradas en otros momentos y puntuadas por otras personas; la tercera, como el indicador del grado en el cual los resultados de una evaluación están libres de los efectos de variación aleatoria causados por factores como la forma de la evaluación, el momento de la aplicación o los calificadores; y, por último, como la consistencia de las mediciones cuando los procedimientos de medición se repiten en una población de individuos o grupos. En este trabajo se calcularon tres índices: el KR-20 a partir de puntuaciones brutas, KR-20 a partir de datos en escala Logit, y la separación. El KR-20 es un caso especial del alfa de Cronbach; en realidad y como lo señalan varios autores, el índice KR-20 indica la repetibilidad de las puntuaciones brutas (que se han malinterpretado como mediciones lineales).
 - e. *Separación*: desde la perspectiva del modelo de Rasch se utiliza el índice de separación (Linacre, 2013) que, desde esta perspectiva, implica o se refiere a la reproductibilidad de las medidas (índice de consistencia interna). Si el valor de la separación es alto (mayor a 2) se considera que el instrumento es sensible para diferenciar el desempeño de las personas.
 - f. *Valores propios*: también conocidos como *eigenvalues*, son valores que se producen con el PCA de residuos –Análisis de Componentes Principales– (Linacre, 2013), y se utilizan con el fin de analizar la dimensionalidad del instrumento (Linacre, 2013). Se espera que el valor propio de la varianza no explicada por el primer contraste sea inferior a 2 (dos ítems) que sería lo mínimo para considerar una dimensión diferente.

Datos obtenidos

Tras el procesamiento de los datos en toda la población evaluada se obtuvieron los diferentes valores para determinar la calidad técnica del instrumento. En la Tabla 10 se presenta un resumen de estos indicadores.

Tabla 10. Resumen de estadísticas del instrumento

N.º de ítems	N.º de evaluados	KR-20	Confiabilidad Rasch	Separación	Promedio	Desviación estándar	Valor propio
30	192	0.96	0.92	3.36	1.4782	1.201	3.7932

De acuerdo con esta información, se considera que la confiabilidad (en los tres datos) presenta valores que indican una alta repetibilidad de la información y que la prueba es lo suficientemente sensible para discriminar el desempeño de los evaluados. El promedio alto que se encuentra en la escala Logit del modelo indica que, en su conjunto, la población evaluada considera que ha desarrollado las habilidades y competencias en alta medida, y son homogéneos en su apreciación (desviación estándar). Por otro lado, el valor propio indica que el primer contraste de ítems tiene la fuerza de casi cuatro ítems; sin embargo, considerando que la prueba se realizó para dar cuenta de seis categorías distintas, se sugiere una calificación separada e individualizada para cada una de esas categorías –a manera de subpruebas–.

Así, después de procesar los datos considerando las seis subpruebas establecidas en el instrumento y al hacer el análisis respectivo, se concluye que en todos los casos los valores indican una alta calidad técnica del instrumento, y que se cumplen los criterios y estándares establecidos (véase Tabla 11).

Tabla 11. Resumen de estadísticas por dimensión del instrumento

Dimensión	N.º de ítems	KR-20	Confiabilidad Rasch	Separación	Promedio	Desviación estándar	Valor propio
1	7	0.90	0.81	2.09	2.30	1.72	2.64
2	4	0.82	0.77	1.81	1.55	1.54	2.07
3	3	0.90	0.87	2.64	3.62	368.38	1.73
4	7	0.91	0.85	2.36	1.92	1.68	2.19
5	4	0.87	0.82	2.15	2.05	1.91	1.66
6	5	0.91	0.85	2.39	2.92	2.32	1.76

En los indicadores para los ítems de la prueba –procesando cada subprueba por separado, como se sugiere en el análisis anterior– se observa que todos los ítems tienen

indicadores dentro los límites establecidos como de buena calidad y ajuste al modelo (véase Tabla 12).

Tabla 12. Indicadores de los ítems

Ítem	Dificultad	Ajuste próximo	Ajuste lejano	rpbis
1a	-0.36	1.16	1.14	0.65
1b	-0.07	0.70	0.68	0.75
1c	-0.40	0.81	0.82	0.71
1d	-0.65	0.99	0.89	0.68
1e	0.25	1.14	1.00	0.72
1f	0.54	1.17	1.20	0.72
1g	0.69	1.19	1.18	0.72
2a	0.08	1.12	1.14	0.71
2b	0.04	0.95	0.93	0.74
2c	-0.07	1.04	1.05	0.72
2d	-0.05	0.89	0.88	0.74
3a	-0.18	1.01	0.89	0.86
3b	0.22	0.77	0.71	0.89
3c	-0.04	1.22	1.17	0.84
4a	-0.11	1.13	1.13	0.70
4b	-0.05	1.08	1.04	0.70
4c	-0.39	0.75	0.75	0.73
4d	-0.14	0.90	0.94	0.73
4e	0.74	1.56	1.66	0.70
4f	0.09	0.71	0.83	0.76
4g	-0.14	0.85	0.84	0.74
5a	0.08	1.18	1.24	0.78
5b	0.55	1.20	1.21	0.79
5c	-0.34	0.73	0.69	0.82
5d	-0.29	0.96	0.89	0.79
6a	-0.21	0.72	0.82	0.79
6b	-0.08	0.91	0.87	0.80
6c	-0.25	0.56	0.62	0.80
6d	0.05	0.88	0.97	0.77
6e	0.49	1.74	1.68	0.74

Resultados del impacto

Los resultados del impacto se presentan de dos maneras diferentes. En primer lugar, se muestran los resultados de la parte inicial, es decir, la primera aplicación del instrumento, en la que se incluyen los estudiantes de séptimo semestre –que actúan como el grupo contrario a los hechos–; se espera que sus resultados sean más bajos que los de los demás estudiantes, y que los de octavo y noveno semestre tengan resultados más altos. Es necesario puntualizar que esta aplicación del instrumento se realizó en la última semana de septiembre, lo que implicó que los estudiantes que empezaban octavo semestre pudieron haber desarrollado un poco las competencias, habilidades y acciones sobre las que se cuestiona. En segundo lugar, se muestra los resultados del segundo proceso, en el que se administró el instrumento al finalizar el semestre; en esta ocasión, no se incluyeron estudiantes de séptimo semestre, ya que se pretendió analizar si hay un cambio en los resultados de los estudiantes luego de un semestre de haber cursado el trabajo de grado y las prácticas (en el caso de los estudiantes de octavo semestre) y al terminar su formación de dos semestres de práctica y de trabajo de grado (los estudiantes de noveno semestre).

Resultados de impacto a partir del primer proceso

En este caso, se administró la prueba y se obtuvieron resultados generales y por dimensión evaluada para cada uno de los 192 estudiantes participantes. Para estimar las diferencias entre los promedios se trabajó con la *t* de Student, utilizando la corrección de Welch (Linacre, 2013), donde se propone un refinamiento para el caso en que las varianzas no sean iguales. El valor de probabilidad (*p*) reportado corresponde a un análisis de dos colas, donde se considera un nivel de significación (es decir, que la diferencia es significativa) si $p < 0.05$ (Linacre, 2013). A continuación se presentan los datos de promedio para cada semestre y alternativa de trabajo de grado y de práctica; en las tablas 13 y 14 se encuentra el valor de *p* para determinar si la diferencia entre promedios de los casos seleccionados es significativa.

Tabla 13. Promedios de las dimensiones del instrumento según semestre y alternativa de trabajo de grado

Semestre	Alternativa	N	Acciones profesionales	Selección de estrategias	Identificación de problemas	Comunicación de resultados	Aplicación de procesos	Nivel de competencia
7		49	1.5	0.8	2.2	1.2	1.1	2.1
8	Curso E. I.	19	2.7	2.3	5.7	2.5	3.2	3.7
8	Investigación	32	2.6	1.7	3.3	2.1	2.1	3.1
8	Social Comunitario	1	2.6	0.1	1.0	0.8	0.9	2.4
8	Movilidad	2	1.1	0.8	6.0	1.9	4.1	2.9
9	Curso E. I.	34	2.6	1.6	4.1	2.1	2.0	2.7
9	Investigación	32	2.3	2.0	4.1	2.2	2.5	3.4
9	Social Comunitario	9	3.2	2.3	5.1	2.7	2.8	4.0
9	Movilidad	4	3.4	2.0	4.0	2.0	1.7	2.0
Total		192	2.3	1.5	3.6	1.9	2.0	2.9

Tabla 14. Promedios de las dimensiones del instrumento según semestre y alternativa de práctica

Semestre	Alternativa	N	Acciones profesionales	Selección de estrategias	Identificación de problemas	Comunicación de resultados	Aplicación de procesos	Nivel de competencia
7		49	1.5	0.8	2.2	1.2	1.1	2.1
8	Clínica	35	2.3	1.4	3.5	1.9	2.1	3.1
8	Organizacional	15	3.3	2.5	5.3	3.0	3.6	4.1
8	Comunitaria	10	2.3	1.9	3.9	2.1	2.4	3.4
9	Clínica	30	2.2	1.9	2.7	1.6	1.8	2.1
9	Organizacional	21	2.4	1.5	4.8	2.6	2.7	3.0
9	Educativa	10	2.8	2.6	6.0	2.3	2.8	4.8
9	Comunitaria	8	3.7	1.9	7.5	3.6	2.9	5.2
9	Jurídica	13	2.8	1.6	3.4	1.7	1.8	3.0
Total		192	2.3	1.5	3.6	1.9	2.0	2.9

En las tablas 13 y 14 el total no corresponde a la suma de los casos de cada semestre y alternativa debido a que algunos estudiantes no marcaron su respuesta correctamente. Sin embargo, estos datos son comparables gracias a que la información se procesó del único cuestionario respondido por los estudiantes. Específicamente, la dimensión de identificación de problemas es en la que los resultados son mayores en ambos casos; en el seminario de trabajo de grado los estudiantes obtuvieron el menor valor en Servicio Social Comunitario, mientras que en la práctica fue en Psicología Clínica. Es importante tener en cuenta que en algunos casos el tamaño de la población es muy pequeño.

Tabla 15. Promedio por semestre

Semestre	Promedio
7	0.09094
8	1.7246
9	1.6347

Siguiendo los planteamientos metodológicos iniciales se estableció si hay diferencia significativa entre estos resultados (véase Tabla 16). Como se puede observar, se concluye que hay diferencias significativas entre el séptimo semestre y el octavo y noveno semestre, pero no la hay entre estos dos últimos; por tanto, una primera conclusión es que tanto el trabajo de grado como la práctica contribuyen realmente a la formación de competencias para el ejercicio de la profesión, lo que implica que solo cuando han empezado los dos últimos semestres de la carrera se inicia el proceso de formación en esas competencias, habilidades y acciones, motivo por el cual los estudiantes de séptimo semestre tienen resultados mucho más bajos.

Tabla 16. Significatividad de las diferencias de promedio según semestre

Semestre 1	Semestre 2	Diferencia	Grados de libertad	p
7	8	-0.815	99	0
7	9	-0.725	128	0
8	9	0.09	122	0.68

Adicionalmente, al establecer promedios y diferencias se trabajó con el software Amos y con la metodología relacionada con las ecuaciones estructurales; se establecieron los pesos de regresión (Arbuckle, 2013), que son una medida de la relación entre las variables; y se presentaron los pesos de regresión estandarizados (sus unidades son desviaciones estándar) y el valor de p para establecer la significatividad de la relación.

A su vez, se produjeron los tamaños de efecto totales (Arbuckle, 2013), definidos como la medida que refleja la cantidad de “efecto” de las dimensiones de la prueba en el resultado total. En el presente caso se trabajaron los efectos estandarizados. Específicamente, las variables que se relacionaron son las dimensiones evaluadas en la prueba y el resultado total de la prueba; y se procesaron los datos para cada alternativa de trabajo de grado y de práctica con el fin de identificar los casos en que la alternativa contribuye a mejorar las competencias, habilidades o acciones evaluadas en el instrumento y que son primordiales para el ejercicio profesional de los futuros profesionales. La expresión gráfica de la relación se muestra en la Figura 9.

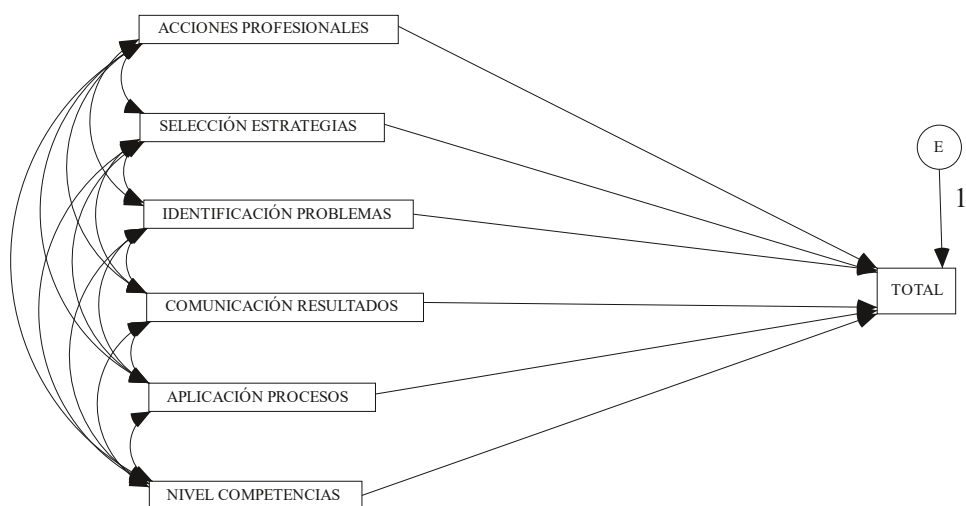


Figura 9. Variables relacionadas con respecto al impacto de las dimensiones en la puntuación total.

Los resultados de los pesos de regresión indican si la relación es significativa. Los asteriscos (*) en la columna denominada *p*, indican que la relación entre las variables es significativa a 0.01 %. Se observa que la única variable que no tiene una relación significativa es la de identificación de problemas. El porcentaje de la varianza explicada por las variables en el total del resultado es de 89 %.

Tabla 17. Pesos de regresión

		Relación	Estimado	P
Total	<---	Acciones profesionales	0.159	***
Total	<---	Selección de estrategias	0.174	***
Total	<---	Identificación de problemas	0.001	0.971
Total	<---	Comunicación de resultados	0.199	***
Total	<---	Aplicación de procesos	0.158	***
Total	<---	Nivel de competencias	0.090	***

Por otra parte, al analizar los datos de la variable trabajo de grado (véase Tabla 18), se encontró que no aparece la alternativa de movilidad por la poca cantidad de datos recogidos, lo que no permitió hacer los análisis necesarios; asimismo, de acuerdo con lo encontrado el impacto del curso de especial interés en el desarrollo de las competencias relacionadas con la identificación de problemas se encontró que las diferencias no son significativas, al igual que en el caso de la variable investigación; en esta última alternativa, el impacto no es significativo en el nivel de competencias ni tampoco en la selección de estrategias; finalmente, en la alternativa social comunitaria se destaca el hecho que el impacto es negativo, es decir, contrario a su desarrollo tanto en la identificación de problemas como en el nivel de competencias.

Tabla 18. Pesos de regresión de las variables analizadas según trabajo de grado

Relación			Curso El		Investigación		Social Comunitaria	
			Valor	P	Valor	P	Valor	P
Total	<---	Acciones profesionales	0.127	0.001	0.221	***	0.71	***
Total	<---	Selección de estrategias	0.2	***	0.099	0.105	0.279	0.096
Total	<---	Identificación de problemas	0.002	0.915	0.009	0.7	-0.096	0.002
Total	<---	Comunicación de resultados	0.121	0.007	0.289	***	0.899	***
Total	<---	Aplicación de procesos	0.152	***	0.172	0.001	0.008	0.94
Total	<---	Nivel de competencias	0.125	***	0.049	0.174	-0.461	***

Adicionalmente, en la Tabla 19 aparecen los pesos de regresión según la práctica realizada. Aquí, es importante destacar el caso de la psicología jurídica, que es la única práctica que tiene un impacto positivo y significativo en todas las variables. Por otro lado, es importante tener en cuenta la contribución o impacto negativo de la práctica en educativa y social comunitaria en algunas de las variables analizadas.

Tabla 19. Pesos de regresión de las variables según la práctica realizada

Relación			Clínica		Organizacional		Educativa		Comunitaria		Jurídica	
			Valor	p	Valor	p	Valor	p	Valor	p	Valor	p
Total	<---	Acciones profesionales	0.214	***	0.08	0.166	0.534	0.184	0.177	0.11	0.071	***
Total	<---	Selección de estrategias	0.109	0.032	0.3	0.002	-0.183	0.444	0.104	0.52	0.151	***
Total	<---	Identificación de problemas	-0.004	0.852	0.028	0.289	0.087	0.378	0.012	0.86	0.019	***
Total	<---	Comunicación de resultados	0.216	***	0.124	0.052	0.234	0.288	-0.205	0.274	0.191	***
Total	<---	Aplicación de procesos	0.134	***	0.166	0.014	0.093	0.812	0.588	0.01	0.092	***
Total	<---	Nivel de competencias	0.082	0.005	0.104	0.013	0.306	0.266	0.224	0.17	0.073	***

Finalmente, al revisar el tamaño del efecto de las variables según cada alternativa de trabajo de grado y cada alternativa de la práctica (véase Tabla 20), lo primero que llama la atención es el efecto negativo de algunas alternativas en algunas variables, lo que implica un análisis profundo y detallado de las alternativas y los sitios en donde se desarrollan, o de la metodología que emplearon. Asimismo, se destacan los altos efectos de algunas alternativas en las acciones profesionales (trabajo de grado, social comunitaria y práctica en educativa); los valores más altos en cuanto al impacto, además de los dos mencionados, son los de práctica social comunitaria en la aplicación de procesos; práctica en organizacional en la selección de estrategias; trabajo de grado en investigación; y social comunitaria en comunicación de resultados; todo esto contrario a lo analizado con los pesos de regresión. En este caso, el tamaño del efecto de la práctica en jurídica es bajo para todas las variables investigadas.

Tabla 20. Tamaño del efecto de las alternativas de trabajo de grado y de práctica en las variables consideradas

	Nivel de competencias	Aplicación de procesos	Comunicación de resultados	Identificación de problemas	Selección de estrategias	Acciones profesionales
Total	0.167	0.236	0.276	0.001	0.216	0.214
Curso E. I.	0.235	0.255	0.199	0.007	0.258	0.187
Investigación	0.092	0.232	0.359	0.024	0.124	0.276
Social Comunitario	-0.665	0.009	0.964	-0.22	0.258	0.688
Movilidad	N/A	N/A	N/A	N/A	N/A	N/A
Clínica	0.176	0.219	0.297	-0.012	0.147	0.297
Organizacional	0.192	0.237	0.174	0.076	0.346	0.118
Educativa	0.269	0.118	0.275	0.17	-0.2	0.504
Comunitaria	0.332	0.69	-0.286	0.026	0.115	0.235
Jurídica	0.148	0.204	0.324	0.068	0.321	0.155

Resultados de impacto a partir del segundo proceso

Lo que se hizo fue comparar los resultados de los estudiantes entre la primera y la segunda aplicación, es decir, la valoración pre y post. Se incluyeron los estudiantes de séptimo semestre como contraste. Como se observa en la Tabla 21, las únicas diferencias significativas aparecen entre los estudiantes de séptimo semestre y la aplicación pre y post. Asimismo, no hay diferencias significativas entre el pre y el post de los otros dos semestres, y una posible explicación es que la aplicación pre en realidad se realizó dos meses después de haber iniciado semestre, lo que puede incidir en un proceso de formación que se refleja en un mayor desarrollo de competencias, habilidades y acciones.

Tabla 21. Diferencia entre las evaluaciones pre y post

Grupo		Diferencia	Grados de libertad	p
7	Pre	-0.8157	133	0.000
7	Post	-1.1047	161	0.000
Pre	Post	-0.289	216	0.125

Conclusiones

El trabajo de grado y la práctica contribuyen al desarrollo de habilidades, competencias y acciones relacionadas con el ejercicio profesional del psicólogo. En su conjunto, estos resultados deben contribuir para un análisis detallado por parte de los docentes y del área de prácticas con el fin de encontrar las posibles causas de los resultados en donde el impacto no es significativo o es negativo. Igualmente, es importante analizar el contexto y los sitios de práctica para ver la contribución de cada uno de ellos en la formación de los estudiantes.

Por otra parte, los resultados aquí reportados, aunque son importantes y significativos, deben ser considerados como preliminares, por lo que sería deseable que se continuara con la evaluación del impacto y que se tenga en cuenta algunas recomendaciones, como las que se presentan a continuación.

Aunque el instrumento mostró adecuadas condiciones psicométricas, sería importante que también fuera validado a través de jueces, con los aportes que puedan dar los docentes de la facultad, ya que son ellos los encargados de favorecer el desarrollo de estas habilidades y competencias en los estudiantes. Adicionalmente, se podría ampliar el número de ítems utilizados, especialmente en algunas dimensiones, lo que podría redundar en la obtención de mayor información. Asimismo, las aplicaciones deben hacerse a más tardar en la segunda y última semana del semestre, de tal manera que los resultados no se vean sesgados por los aprendizajes adquiridos durante el semestre.

Adicionalmente, la evaluación de impacto debe ir acompañada de un análisis de las condiciones y contribuciones que los diferentes sitios de práctica facilitan a los estudiantes; y sería importante, también, contar con un número mayor y más significativo de estudiantes en las distintas modalidades de trabajo de grado y práctica, de manera que se puedan hacer consideraciones para cada una de ellas (aspecto que no se logró en estos resultados, especialmente porque en la modalidad de movilidad la cantidad de estudiantes era muy baja).

También, aunque la presente investigación fue de corte transversal, el equipo de trabajo considera que, ya que se cuenta con los resultados pre de los estudiantes de séptimo semestre, es posible y recomendable continuar con una investigación de corte longitudinal, donde se evalúe semestre a semestre el desarrollo de las competencias; esto sentaría las bases para tener información más precisa sobre la evaluación del impacto tanto de las prácticas profesionales como del trabajo de grado.

Finalmente, sería interesante incluir otros programas (especializaciones, por ejemplo) en los que los estudiantes también cursan estas dos asignaturas, de manera que la Facultad cuente con información suficiente y precisa sobre estos dos temas, que son de gran relevancia, especialmente si se tiene en cuenta que en el panorama actual se está desarrollando el Plan 8.

Referencias

- Aiken, L. (1996). *Rating scales & checklist. Evaluating behavior, personality, and attitude*. New York: John Wiley & Sons, inc.
- Aldercotte, A. (2018). *Monitoring and evaluating impact*. Reino Unido: Equality Challenge Unit. Recuperado de http://www.ecu.ac.uk/wp-content/uploads/2018/03/2018-03-ECU_Research-and-data-briefing-5_final.pdf
- Arbuckle, J. (2013). *IBM® SPSS® Amos™ 22 User's Guide*. Armonk, Estados Unidos: IBM.
- Chong, H. Y. (2010). A Simple Guide to the Item Response Theory (IRT) and Rasch Modeling. Recuperado de [http://www.fisica.net/enem/A-Simple-Guide-to-the-Item-Response-Theory-\(IRT\)-and-Rasch-Modeling.pdf](http://www.fisica.net/enem/A-Simple-Guide-to-the-Item-Response-Theory-(IRT)-and-Rasch-Modeling.pdf)
- Equality Challenge Unit. (s. f.). *Conducting Equality Impact Assessments in Higher Education*. Londres.
- Gertler, P., Martinez, S., Premand, P., Rawlings, L., & Vermeersch, C. (2007). *Impact evaluation in practice*. Washington, D. C.: World Bank.
- Independent Evaluation Group, IEG. (2012). *World Bank group Impact Evaluation. Relevance and Effectiveness*. Washington, D. C.: IEG.
- Jones, L., & Thissen, D. (2007). A history and overview of psychometrics. En *Handbook of Statistics*. North Holland: Elsevier.
- Kandkher, S., Koolwal, G., & Samal, H. (2010). *Handbook on Impact Evaluation. Quantitative Methods and Practices*. Washington, D. C.: The World Bank.
- Legovini, A., Di Maro, V., & Piza, C. (2015). *Impact Evaluation Helps Deliver Development Projects*. Washington, D. C.: World Bank, Development Research Group.
- Linacre, J. M. (2013). Winsteps® (Version 3.80.0) [Software de computador]. Beaverton, Oregon: Winsteps.com. Recuperado de <http://www.winsteps.com/>
- Mislevy, R., Almond, R., & Lukas, J. (2003). *A brief introduction to evidence Centered Design. Research report*. Princeton, N. J.: Educational Testing Service.
- Misra, V. (s. f.). *Social Impact Assessment Methodology*. Centre for Good Governance.
- Pardo, C. (2010). Manual para el procesamiento y análisis de datos (capítulos 7 y 8). En UNESCO/OREALC, *Compendio de los manuales del SERCE*. Santiago de Chile: Unesco.
- Pardo, C. & Rocha, M. (2010). *Reglas para elaborar ítems de formato de selección y de producción*. Santiago de Chile: UNESCO/OREALC.