



Revista electrónica de investigación educativa

ISSN: 1607-4041

Universidad Autónoma de Baja California, Instituto de Investigación y Desarrollo Educativo

Sancho-Álvarez, Carlos; Jornet Meliá, Jesús Miguel; González-Such, José
Propiedades métricas de un instrumento para evaluar el Valor Social Subjetivo
de la Educación: escala SECS-EVALNEC VSE-Estudiantes-Secundaria1

Revista electrónica de investigación educativa, vol.

20, núm. 3, Julio-Septiembre, 2018, pp. 92-104

Universidad Autónoma de Baja California, Instituto de Investigación y Desarrollo Educativo

DOI: 10.24320/redie.2018.20.3.1762

Disponible en: <http://www.redalyc.org/articulo.oa?id=15557366009>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

UAEM  redalyc.org

Sistema de Información Científica Redalyc

Red de Revistas Científicas de América Latina y el Caribe, España y Portugal

Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

Vol. 20, Núm. 3, 2018

Propiedades métricas de un instrumento para evaluar el Valor Social Subjetivo de la Educación: escala SECS-EVALNEC VSE-Estudiantes-Secundaria¹

Metric Properties of an Instrument to Assess the Subjective Social Value of Education: SECS-EVALNEC SVE-Students-Secondary Education Scale

Carlos Sancho-Álvarez (*) carlos.sancho@uv.es

Jesús Miguel Jornet Meliá (*) jornet@uv.es

José González-Such (*) jose.gonzalez@uv.es

*Universidad de Valencia

(Recibido: 21 de diciembre de 2016; Aceptado para su publicación: 6 de marzo de 2017)

Cómo citar: Sancho-Álvarez, C., Jornet, J. M. y González-Such, J. (2018). Propiedades métricas de un instrumento para evaluar el Valor Social Subjetivo de la Educación: escala SECS-EVALNEC VSE-Estudiantes-Secundaria. *Revista Electrónica de Investigación Educativa*, 20(3), 92-104. Recuperado de <https://doi.org/10.24320/redie.2018.20.3.1762>

Resumen

El estudio analiza los resultados de la aplicación de un instrumento que pretende medir el Valor Social Subjetivo de la Educación. Trabajando con un grupo de 341 estudiantes de Educación Secundaria Obligatoria en España, se exponen las propiedades métricas del instrumento desde una metodología de Teoría Clásica de los Tests y Teoría de Respuesta al Ítem para comprobar su fiabilidad y bondad de ajuste. Después de comprobar la dimensionalidad de la escala y el funcionamiento de los reactivos, los datos muestran que todos se ajustan adecuadamente al modelo de referencia. A partir de ello, presentamos una propuesta de escala SECS-EVALNEC VSE-Estudiantes-Secundaria unidimensional conformada por 20 ítems con una fiabilidad global de .73. Los resultados del estudio demuestran niveles aceptables en ambas perspectivas, aportando un primer indicio como evidencia de validación del instrumento de evaluación.

Palabras clave: Evaluación educativa, valor social de la educación, escala de evaluación, validación, Teoría de Respuesta al Ítem.

Abstract

The study analyzes the results of the administration of an instrument intended to measure the subjective social value of education. By working with a group of 341 students of compulsory secondary education in Spain, the study presents the metric properties of the instrument based on a classical test theory and

¹ En el marco del proyecto I+D+I "Sistema Educativo y Cohesión Social: diseño de un modelo de evaluación de necesidades (SECS/EVALNEC)", ref. EDU2012-34734 financiado por el Ministerio de Economía y Competitividad. Así como con "Ajudes per a la formació de personal investigador de caràcter predoctoral, en el marc del Subprograma Atracció de Talent 2013" del Vicerectorat d'Investigació i Política Científica de la Universitat de València (España).

item response theory methodology to verify its reliability and goodness-of-fit. After the dimensionality of the scale and the functioning of items were verified, all items displayed an appropriate fit to the benchmark model. On this basis, a proposal is put forward for a unidimensional SECS-EVALNEC VSE-*Students-Secondary Education* scale made up of 20 items with a global reliability of .73. The results of the study show acceptable levels in both perspectives, providing initial evidence to validate this assessment instrument.

Keywords: Educational assessment, social value of education, evaluation questionnaire, test validity, Item Response Theory.

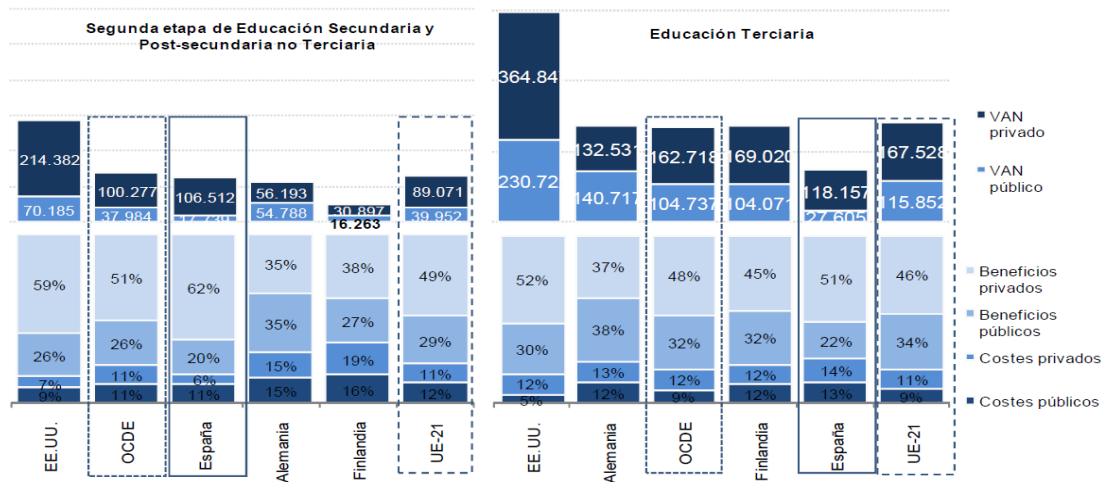
I. Introducción

La idea de que la educación genera beneficios sociales es cada vez más generalizada. En las últimas décadas diversos estudios han señalado de forma consistente la correlación entre niveles elevados de educación y mejores condiciones de salud en los individuos. Una posible hipótesis para explicar estos resultados puede partir del supuesto de que las personas con diferentes niveles educativos poseen distinto conocimiento sobre cómo llevar conductas saludables. Por lo tanto, los individuos con mayores niveles educativos, teóricamente mejor informados, escogerían comportamientos más saludables (Ministerio de Educación, Cultura y Deporte, 2013, p. 27).

En el ámbito de la salud es frecuente el estudio objetivo de la repercusión que tiene la educación sobre la misma, quedando clara su influencia directa. Sin embargo, en el sector educativo todavía es un tema a explorar, además de que podría ser más complejo, variable y dependiente del contexto.

La educación es, sin duda, un factor clave de promoción y desarrollo personal y transformación social. Sin embargo, resulta llamativo que desde los estudios de Coleman et al. (1966) siempre se identifique el nivel socioeconómico y cultural como el predictor más claro del nivel de logro. Junto a ello, cuestiones como el esfuerzo, la motivación hacia el estudio, la implicación del propio estudiante en su aprendizaje, los apoyos y expectativas familiares, se apuntan en diferentes trabajos como variables relacionadas con el logro.

El hecho de que la educación tenga beneficios para la mejora del estatus personal es un factor clave de prevención de la exclusión social; en la actualidad, entre los expertos nadie lo duda. Los estudios derivados de evaluaciones internacionales así lo evidencian de forma sistemática. En esta línea argumentativa, cabe resaltar el caso de algunos indicadores que representan esta realidad. Ya que en todos los países de la OECD, según datos del 2012 (OECD, 2013), los sujetos que poseen un nivel de educación superior tienen más probabilidad de un empleo con jornada completa y las tasas de empleo son mucho más altas que en los casos que poseen un nivel educativo inferior. Asimismo, las tasas de empleo son mayores entre las personas que poseen un nivel de educación secundaria superior en comparación con los que no la tienen. También, de manera objetiva, se ha podido constatar la influencia de la educación sobre algunos aspectos de rentabilidad económica. De acuerdo con el Ministerio de Educación, Cultura y Deporte [MECD] (2013), respecto a la inversión realizada por todos los países de la OECD en educación, se superan las ganancias esperadas en todos los casos, siendo tanto la inversión pública como la privada altamente rentables. De manera exponencial, se observa su incremento en el aumento del nivel educativo sobre los beneficios absolutos, que pueden ser evaluados mediante el Valor Actual Neto (VAN) (ver figura 1).



Fuente: OECD (2013, p. 28).

Figura 1. Datos sobre Valor Actual Neto entre los países de la OECD

Por lo tanto, sí que se observan tendencias hacia la explicación del Valor Social Objetivo de la Educación de manera aislada, pero habría que profundizar en aspectos subjetivos de este constructo de manera integrada. Para ello, se identifican algunos modelos de evaluación educativa concretos que trabajan de manera sistémica el proceso educativo (De la Orden, 2007). Jornet, Sánchez-Delgado y Perales (2015) plantean un modelo para evaluar el impacto y la relevancia de la educación en la sociedad. Jornet (2012) propone otro modelo, donde considera la evaluación como un facilitador esencial para caminar hacia la cohesión social desde el ámbito educativo. En este caso, éste será el marco de referencia para el presente trabajo, ya que encontramos aspectos evaluativos relevantes que, junto con características de contexto, se proponen diversos constructos teóricos socio-educativos donde el del Valor Social de la Educación será el principal foco de interés a partir de ahora.

De esta manera, se podrán vislumbrar horizontes que ayuden a progresar en la práctica docente y en los resultados del alumnado, así como a encontrar pautas de mejora de la práctica escolar, todo ello desde la evaluación de sistemas educativos (Jornet, López-González y Tourón, 2012). Algo siempre necesario será el indiscutible análisis de los aspectos evaluativos desde un enfoque contextual del aprendizaje, para conocer los factores asociados e incluso influyentes sobre la educación (González-Montesinos y Backhoff, 2010; Joaristi, Lizasoain y Gamboa, 2012; González-Barbera, Caso-Niebla, Díaz-López y López-Ortega, 2012).

No obstante, el convencimiento de que la educación puede actuar como un “ascensor social” –como la denominan autores como Valcárcel (2013)– entre la población constituye una base clara que apoya el incremento de la motivación y favorece el esfuerzo e implicación del alumnado. Evaluar este aspecto es, pues, necesario.

Tal como se expone en trabajos de Jornet, Perales y Sánchez-Delgado (2011), en el marco del Modelo de Evaluación de Sistemas basados en el concepto de Cohesión Social (Jornet, 2012) se propone considerar el Valor Social que las personas (fundamentalmente el alumnado) dan a la educación, como un constructo que, de forma holística, permita acercar la identificación de elementos socio-afectivos que subyacen sobre la motivación hacia el estudio. Por este motivo, se diseña una escala de evaluación que pueda permitir valorar el grado en que el alumnado cree que la educación es un factor importante para su promoción y desarrollo personal –considerados estos elementos como el Valor Social Subjetivo de la Educación (vse-Subjetivo). Para ello se llevan a cabo diversos estudios encaminados a identificar evidencias de validez de constructo basadas en comités de expertos. Para tal efecto, partimos del constructo teórico validado por Sancho-Álvarez, Jornet y González-Such (2016), así como por la propuesta de escalas de evaluación sobre el vse-Subjetivo en estudiantes (Sancho-Álvarez, Jornet y González-Such, 2017).

En este trabajo presentamos el análisis de las propiedades métricas de la escala desarrollada. Nos basamos en un ensayo piloto aplicado en España y cuyas características se comentan posteriormente.

II. Método

Se trata de un estudio de diseño y validación de instrumentos de medición de carácter socio-afectivo, basado en las percepciones del alumnado. Se sustenta en dos aproximaciones: Teoría Clásica de los Tests (TCT) y Teoría de Respuesta al Ítem (TRI), modelo de un parámetro (Rasch) (Delgado, 2014; González-Montesinos, 2008; Linacre y Wright, 2009). El carácter del mismo es el de ensayo piloto empírico, cuyo objetivo principal es depurar la escala, ajustándola para conseguir los mejores niveles en cuanto a sus propiedades métricas. Los objetivos específicos son: a) Analizar la calidad de los reactivos en cuanto a los diferentes indicadores que proveen los modelos TCT y TRI, b) Identificar los ítems defectuosos para proceder a la depuración de la escala, y c) Valorar la fiabilidad global de la escala depurada.

Se trabajó con un grupo, no con una muestra estadísticamente representativa. Para la incorporación de participantes se realizaron diversos cursos de formación dirigidos a equipos directivos de centros educativos de la Comunidad Valenciana (Alicante, Castellón y Valencia) orientados a la evaluación de centros basados en el Modelo de Evaluación desde el concepto de Cohesión Social (como se citó en Jornet, 2012). Por ello, la participación del alumnado encuestado ha sido por voluntariedad del centro para colaborar en la experiencia completa de evaluación institucional.

Se repartieron cuestionarios en formato papel y se facilitó también una plataforma en línea, dejando a elección el formato de cumplimentación. Se completaron en castellano un 51.6% y en valenciano el 48.4%². Así como un 94.4% en lápiz y papel, siendo el resto en formato electrónico (5.6%). Por último, al tomar como criterio de exclusión las respuestas incompletas, el grupo de estudio quedó formado por 341 estudiantes de Educación Secundaria Obligatoria (ESO), 49% chicas y 51% chicos.

Los análisis³ de propiedades métricas realizados a partir de los resultados de la aplicación del instrumento de evaluación del VSE-Subjetivo se orientaron tanto desde el enfoque de la TCT como desde el modelo Rasch sobre la TRI.

- Comprobación de supuestos de normalidad de la distribución y unidimensionalidad de la escala. Para ello se incluyeron: Estadísticos descriptivos y frecuencias para observar datos generales, Prueba Kolmogorov-Smirnoff para comprobar la normalidad de la distribución y exploración dimensional a partir de una estrategia basada en Análisis de Componentes Principales (con rotación oblicua), con exploración de soluciones de primer, segundo y tercer orden.
- Como análisis métricos hemos utilizado dos aproximaciones: TCT (modelo α o de consistencia interna de Cronbach), y TRI (un parámetro, modelo de Rasch). La utilización de ambas aproximaciones se basa en el hecho de que la depuración de la escala requiere de la máxima información posible que permita comprender el funcionamiento real de los reactivos. Se entiende, pues, que en un proceso de ensayo piloto orientado a la depuración de la escala, necesariamente son complementarias.

En síntesis, los análisis realizados han sido: a) basados en la TCT: pruebas de tendencia central, distribución y variación, análisis de fiabilidad mediante alfa de Cronbach, y depuración técnica de ítems defectuosos; b) basados en el modelo Rasch (TRI): estudio de ajuste estadístico al modelo Rasch mediante el cálculo de INFIT, OUTFIT, MNSQ, ZTSTD, RMSE y otros estadísticos de ajuste; estudio de probabilidades de las

² Aunque no se realizó un estudio sistemático de validación cultural para analizar la equivalencia de la escala en ambos idiomas, se sometió a la consideración de un grupo de expertos bilingüe (castellano/valenciano) si realmente en las dos versiones podían existir problemas vinculados a la traducción o posibles fuentes de sesgo. Téngase en cuenta que ambos idiomas son co-oficiales en la Comunidad Valenciana y que el profesorado está obligado a tener un nivel suficiente de dominio en los dos. Por lo que entendimos que esta opción aportaba suficientes garantías.

³ Los paquetes estadísticos utilizados han sido el SPSS 22 para TCT y WINSTEPS 3.73 para TRI.

categorías de respuestas, estudio de puntuaciones de ítems y estudio de información que aporta la escala (variable latente).

III. Resultados

3.1. Comprobaciones previas de supuestos

Análisis descriptivos. Los resultados se presentan en primer lugar en relación a sus respectivas medidas de tendencia central, dispersión y forma, estructurados por sus dimensiones teóricas sobre la escala y totales⁴ (ver tabla I).

Tabla I. Estadísticos descriptivos por dimensión y totales

DIMENSIONES E ÍTEMS	\bar{x}	σ	CV*	Asim.	Curt.
Dimensión 1. Expectativas y metas educativas					
1_1.1 ¿Tus profesores/as creen que vas a suspender?	2.97	0.97	32.73	-0.63	-0.60
2_1.2 ¿Tus padres quieren que trabajes porque estudiar no sirve para tener más dinero?	3.41	1.00	29.35	-1.46	0.67
3_1.3 ¿Tus padres creen que es más importante estudiar que ganar dinero?	3.07	0.96	31.24	-0.68	-0.60
4_1.4 ¿Tus padres creen que vas a aprobar?	3.08	0.94	30.62	-0.60	-0.77
td1	3.13	0.58	18.64	-0.40	-0.19
Dimensión 2. Justicia Social y Educación					
5_2.1 ¿Es importante para ti ir bien en tus estudios?	3.42	0.93	27.16	-1.53	1.18
6_2.2 ¿En tu centro escolar te ayudan a tener confianza en ti mismo/a?	2.69	0.97	35.91	-0.19	-0.95
7_2.3 ¿En tu centro escolar te enseñan cosas útiles para un futuro trabajo?	2.83	0.91	32.01	-0.35	-0.68
8_2.4 ¿En tu centro escolar te ayudan a encontrar amigos/as?	2.39	1.01	42.34	0.04	-1.11
9_2.5 ¿En tu centro escolar te ayudan a aprender a vivir con los/as demás?	2.73	0.96	35.09	-0.29	-0.85
10_2.6 ¿Lo que estás aprendiendo en tu centro escolar te servirá cuando decidas buscar trabajo?	2.88	0.90	31.28	-0.34	-0.75
td2	2.82	0.61	21.71	-0.02	-0.48
Dimensión 3. Valor diferencial de la Educación					
11_3.1 ¿Estudiar te ayuda a aprobar?	3.35	0.93	27.67	-1.26	0.51
12_3.2 ¿Tienes compañeros/as que aprueban sin merecérselo?	2.32	0.97	41.77	-0.00	-1.09
13_3.3 ¿Piensas que las personas que estudian triunfan más?	3.13	0.87	27.80	-0.75	-0.22
14_3.4 ¿Las personas que triunfan en la vida no han estudiado?	2.85	0.88	30.74	-0.43	-0.47
td3	2.80	0.38	13.67	-0.09	0.34
Dimensión 4. Obstáculos y facilitadores					
15_4.1 ¿Tus padres quieren que trabajes cuanto antes porque hace falta dinero en casa?	3.54	0.83	23.50	-1.73	1.94
16_4.2 ¿Cuanta más gente haya estudiado en tu ciudad mejor para todos/as?	2.81	0.94	33.31	-0.28	-0.86
17_4.3 ¿Ser profesor/a es una profesión muy importante?	3.05	0.82	26.89	-0.43	-0.59
18_4.4 ¿Tus amigos y amigas no quieren estudiar porque piensan que no sirve para nada?	2.92	0.96	32.81	-0.59	-0.56
19_4.5 ¿Los/as políticos/as de tu país hacen muchas cosas para que tu educación sea mejor?	1.91	0.85	44.66	0.59	-0.45
20_4.6 ¿Las personas famosas normalmente no han estudiado nada?	2.58	0.80	31.16	-0.40	-0.33
td4	2.91	0.58	19.82	-0.66	0.41
t	2.89	0.38	13.31	-0.38	0.93

*Cociente de variación

⁴ Para ello, los valores de los ítems formulados en negativo fueron invertidos para adecuarlos a los análisis de datos.

Normalidad. Los puntajes de la escala de valoración oscilan entre 1 y 4 (siendo el 1 la puntuación más baja). Los análisis realizados muestran que la forma de la distribución es asimétrica negativa, concentrándose las puntuaciones en los valores superiores al promedio en cada ítem de la escala. Las distribuciones, además, son bastante homogéneas por dimensiones y sobre la distribución de la puntuación total (CV: 13.31%), lo que indica que el nivel de homogeneidad en las respuestas a cada ítem es bastante elevado.

Se realizó la prueba de normalidad Kolmogorov-Smirnoff y se halló que el grupo de estudio obtiene una distribución no normal, con una significación estadística de 0.005. De este modo, tanto desde análisis paramétricos como no-paramétricos se observaron desajustes en cuanto al supuesto de normalidad.

Dimensionalidad mediante análisis factorial. A partir de obtener un KMO de .754 y una significación de $p \leq .000$ en la prueba de Bartlett se consideran niveles adecuados para la estimación hacia una solución factorial. Se realizó un análisis factorial con rotación oblicua con el fin de explorar la matriz de estructura dimensional. En un primer orden se observan seis factores que explican el 59.5% de la varianza. Para corroborar la estructura de factores se realizó otro análisis factorial de segundo orden, donde se obtienen dos factores que explican el 42.7% de la varianza. Ya desde la primera solución factorial se observa que, en general, las comunialidades apuntan hacia una estructura unidimensional (dado que son elevadas, y muestran que los reactivos participan en conjunto de todos los factores), por lo que se realizó un análisis factorial de tercer orden para descartar o comprobar este indicio. Por último, a partir del tercer análisis no se confirmó la estructura de dos factores y se corroboró la unidimensionalidad de la escala, obteniendo un único factor que explica el 51.37% de la varianza.

Por lo tanto, a la vista de los resultados, se encontró que teóricamente se proponen cuatro dimensiones y empíricamente se encontró un único factor. En este sentido, haber hallado empíricamente un único factor, se expondrán los análisis de propiedades métricas para el total de la escala considerada unidimensional. Se estima que la dimensionalización de contenido, si bien tiene sentido para describir perfiles de respuesta para establecer el puntaje de VSE-Subjetivo, lo adecuado es tratarlo como una Escala Unidimensional.

3.2. Análisis métrico desde el Modelo Clásico del Test (TCT)

Análisis de fiabilidad. El alfa de Cronbach sobre la escala conformado por los 20 ítems es de .728 ($\bar{x} = 2.90$; $\sigma^2 = .166$). A continuación se realiza un resumen de los elementos que restan fiabilidad, con versiones consecutivas, de acuerdo a los ítems (tabla II).

Tabla II. Resumen de reducción de elementos mediante análisis de fiabilidad para total

Versión	N ítems	\bar{x}	σ^2	α	Ítems que restan fiabilidad
0	20	2.90	.166	.728	2.4; 3.2; 4.4; 4.5; 4.6
1	15	3.06	.074	.779	1.1
2	14	3.06	.079	.781	2.5
3	13	3.09	.076	.783	

El estudio de fiabilidad aporta información fundamental sobre las propiedades métricas de la escala y la eliminación de elementos, lo cual permite descender al nivel de análisis de ítems, y valorar la conveniencia de eliminar algunos en aras de mejorar el funcionamiento de la escala. A partir de la relación anterior se adopta como primer criterio, para poder eliminar algún ítem defectuoso, que su supresión no afecte a la validez de contenido. Si cumple este criterio, el segundo será que su supresión no afecte a la fiabilidad global de la escala. Así, se decidió no eliminar ningún ítem, ya que la supresión de cada uno de ellos, anteriormente expuestos como defectuosos, afectaría a la validez de contenido de la escala. Por lo que se realizaron los siguientes análisis con el conjunto de 20 ítems que conforman la escala para analizarlos desde el enfoque de TRI.

3.3. Análisis métrico basado en el Modelo Rasch (TRI)

La validación con TRI se centra en el análisis de la escala como un único factor, ya que este tipo de aproximación parte del respeto a los supuestos básicos de unidimensionalidad e independencia local (González-Montesinos, 2008; Martínez-Árias, Hernández-Lloreda y Hernández-Lloreda, 2006). Por ello, los resultados se estructuran en torno al total de ítems como escala; con estadísticos sumarios, curva característica de ítems, estudio de información que aporta la escala y otros estadísticos de ajuste.

Resultados escala total. Obtenemos un *person raw score to measure correlation* de .98. Así como una media de INFIT de 1.01 y un OUTFIT de .99. Además de un acuerdo entre las personas de .73 con un adecuado nivel de error del .03.

Tabla III. INPUT 20 items measure order total scale

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PT-MEASURE CORR.	EXP.	EXACT OBS%	MATCH EXP%	ITEM	G
19	645	341	1.35	.07	1.31	4.0	1.32	4.0	.12	.42	41.3	42.3	ES268_4.5	0
12	785	341	.82	.06	1.20	3.1	1.23	3.5	.25	.43	34.6	36.8	ES253_3.2	0
8	812	341	.63	.06	1.18	2.9	1.22	3.2	.29	.45	30.8	37.0	ES260_2.4	0
20	881	341	.52	.07	1.21	2.8	1.23	2.9	.13	.37	43.1	49.5	ES269_4.6	0
6	918	341	.19	.06	.85	-2.5	.84	-2.6	.57	.42	41.9	40.1	ES258_2.2	0
9	930	341	.18	.06	.99	-.1	.98	-.3	.44	.42	42.5	41.0	ES261_2.5	0
16	954	341	.03	.06	1.04	.6	1.03	.5	.39	.41	41.1	40.8	ES264_4.2	0
7	966	341	.01	.06	.81	-2.9	.80	-3.1	.59	.40	51.6	43.6	ES259_2.3	0
14	975	341	-.01	.07	1.04	.6	1.07	1.0	.33	.39	46.0	46.3	ES266_3.4	0
18	999	341	-.01	.06	1.23	3.1	1.27	3.4	.20	.41	40.5	42.7	ES267_4.4	0
1	1011	340	-.07	.06	1.05	.8	1.05	.7	.36	.41	40.9	41.5	ES251_1.1	0
10	980	341	-.10	.07	.84	-2.5	.83	-2.6	.56	.40	47.2	42.1	ES262_2.6	0
3	1047	341	-.24	.06	1.08	1.2	1.21	2.6	.33	.40	41.6	39.5	ES255_1.3	0
4	1050	341	-.34	.06	.90	-1.6	.86	-1.9	.50	.40	41.6	38.3	ES257_1.4	0
13	1071	341	-.38	.07	.92	-1.1	.90	-1.3	.46	.37	47.5	45.4	ES263_3.3	0
2	1162	341	-.41	.06	.96	-.4	.88	-.8	.43	.38	58.4	51.7	ES254_1.2	0
11	1143	341	-.46	.06	.82	-2.3	.70	-3.0	.53	.37	56.3	48.4	ES250_3.1	0
17	1040	341	-.47	.07	.90	-1.5	.89	-1.6	.48	.37	48.1	45.5	ES265_4.3	0
5	1172	341	-.49	.06	.82	-2.0	.64	-3.1	.52	.36	57.5	53.2	ES252_2.1	0
15	1211	341	-.73	.07	.93	-.6	.79	-1.5	.41	.33	63.0	57.3	ES256_4.1	0
MEAN	987.6	340.9	.00	.06	1.00	.1	.99	.0			45.8	44.2		
S.D.	135.4	.2	.50	.00	.15	2.1	.20	2.4			7.9	5.4		

En cuanto a estadísticos por ítem, en la tabla III se puede comprobar cómo todos los ítems se ajustan en el rango⁵ entre .80 y 1.3 para INFIT y OUTFIT (Boone, Staver y Yale, 2014). Por lo que al obtener una situación productiva óptima en ambos criterios para el proceso de medición en todos los reactivos, se decide no eliminar ninguno.

Asimismo, en las curvas de ajuste no se encuentran graves anomalías fuera de los límites de confianza, ya que coinciden con las expectativas del modelo Rasch, así como también las curvas de probabilidad de respuesta para cada ítem; como por ejemplo se puede ver para el ítem 16_4.2 ¿Cuanta más gente haya estudiado en tu ciudad mejor para todos/as?, un buen comportamiento y ajuste en todos los criterios sobre el modelo de referencia (figura 2).

⁵ Nótese que el criterio tomado para valorar el ajuste es el más restrictivo, pudiendo haber tomado aquel en el que (por la tipología de la variable medida: actitudinal, socio-afectiva, etc.) el rango de aceptación podría haber sido más amplio (Boone, Staver y Yale, 2014).

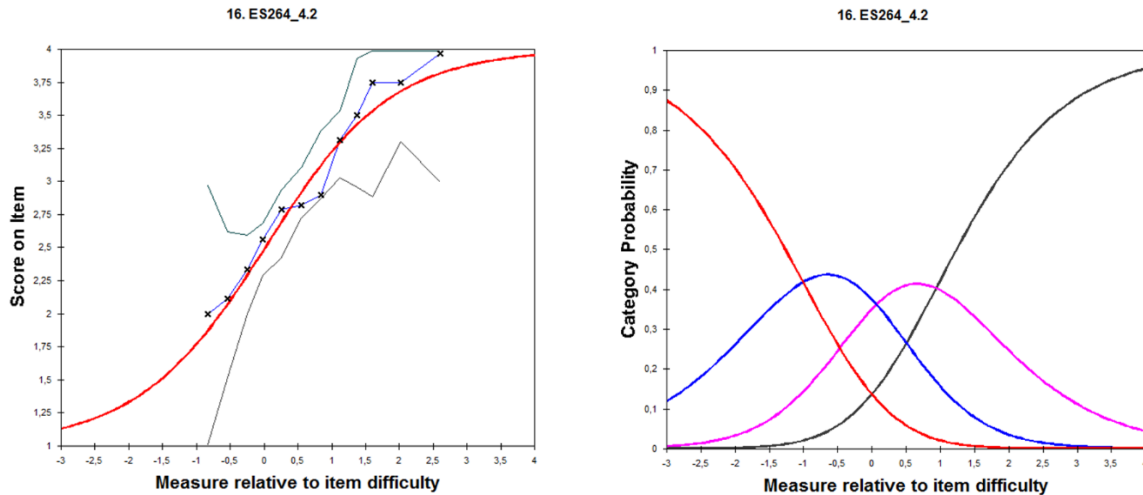


Figura 2. Curva de ajuste y probabilidad de respuesta para el reactivo 16_4.2

El mapa de Wright (figura 3), señala que la media de dificultad coincide para poder responder de forma adecuada en estas preguntas, en cuanto a la media de actitud de los sujetos, ya que la desviación no es dispar. Asimismo, se puede observar que en su mayoría los ítems se comportan de manera distribuida entre todos los niveles, para este grupo de sujetos. Por ello, se ajustan durante toda la curva de actitud discriminando adecuadamente el vSE-Subjetivo, así como las probabilidades de respuesta, que se observan adecuadas para cada reactivo en todos los casos. Sin embargo, sí se observa una mayor discriminación en los niveles altos y medios que en los bajos, ya que existe una diferencia de una desviación entre la media de los ítems y los sujetos. Algo que se debe considerar para futuras propuestas de mejora sobre la formulación de los ítems o su direccionalidad, ya que puede afectar a su nivel de dificultad y sobre la tendencia a contestar valores cercanos a puntuaciones altas.

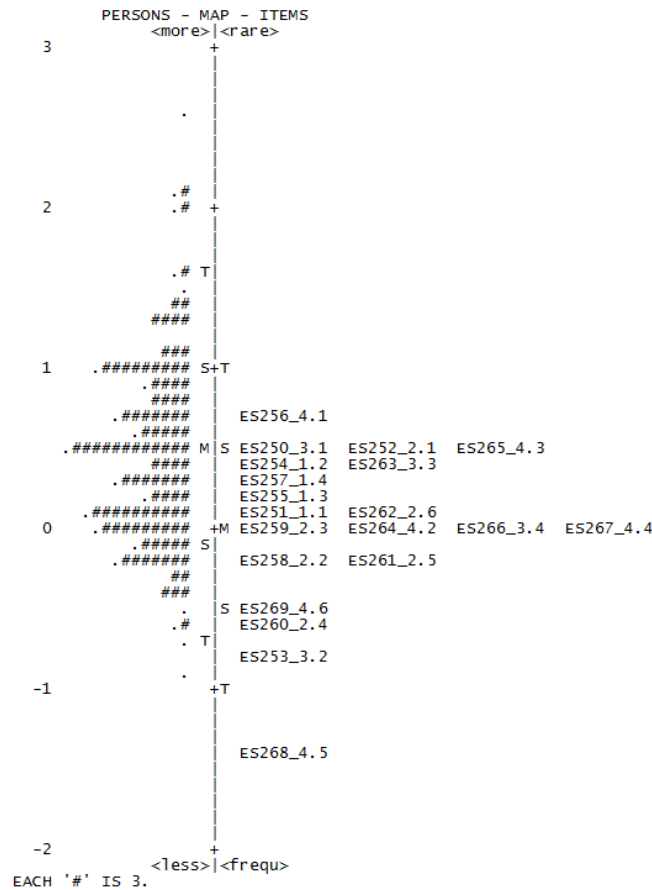


Figura 3. Mapa de Wright Personas e Ítems

En cuanto a la información que recoge la escala completa sobre la curva de información, la mayor información se recoge entre los niveles de -2 a 2 DT de la actitud medida; fuera de ese rango el nivel de información recogida baja drásticamente. Asimismo, la curva característica del test muestra un nivel de actitud medio que corresponde a la puntuación de 50 puntos aproximadamente, algo muy adecuado para los niveles globales de la escala.

IV. Discusión de resultados y conclusiones

En relación con los datos sobre los análisis realizados de propiedades métricas, teniendo en cuenta los dos enfoques metodológicos desarrollados (TCT y TRI), podemos concluir una serie de decisiones para mejorar la fiabilidad y validez de la propia escala:

- A partir de los diferentes análisis de fiabilidad de TCT realizados, se concluye que no se puede eliminar ningún ítem porque no cumplen con los criterios establecidos para su adecuada supresión. El hecho de que a través del análisis realizado a partir de la TCT se identifiquen algunos reactivos cuya eliminación mejoraría ligeramente la fiabilidad total, estimamos que no es indicador suficiente para proceder a su eliminación. Como se señaló antes, la escala –pese a ser unidimensional, lo que implica que las personas sean calificadas por el puntaje total–, contiene dimensiones teóricas que permiten matizar el perfil personal ante sus respuestas a los ítems y dimensiones. Esta virtualidad que se da en muchas escalas y tests estandarizados, lleva a contemplar como prioritaria la validez de contenido, de forma que es preferible (desde nuestro punto de vista) mantenerlos a suprimirlos.

Por otra parte, como se puede observar a partir de los análisis realizados con el modelo de Rasch, todos los ítems mantienen un escalamiento óptimo, lo que constituye un argumento adicional para el mantenimiento de los mismos. Por último, en esta misma línea, hay que señalar que entendemos que la escala, al estar en proceso de diseño, será analizada también con otras muestras, con carácter diferencial. Tal como se indicó, este constructo tiene un carácter evolutivo que se manifestará en diferencias funcionales a través de la edad, por lo que mantener (cuando se analiza) un solo tramo de edad es la solución más adecuada dado que no presentan problemas importantes.

- A partir de los diferentes análisis realizados de TRI se puede concluir que la mayoría de los ítems se ajustan a la curva de expectativa y probabilidad, aunque debería de aplicarse el instrumento en otro grupo de referencia o modificar los ítems para ajustar el nivel de discriminación en los niveles bajos. Este hecho, también identificado por diversos autores como “efecto suelo”, debe corregirse para incrementar la capacidad de discriminación de la escala entre sujetos con bajo VSE-Subjetivo.
- Los dos conjuntos de análisis realizados (TCT y TRI) apoyan en cierta manera no eliminar ningún reactivo, pues no aportan indicios suficientes y congruentes entre ambas aproximaciones para apoyar la decisión de supresión o cambio de ningún elemento, por lo que se aconseja continuar con la línea marcada y apoyarse en posteriores análisis diferenciales de la escala para detectar anomalías.

En definitiva, como se ha constatado mediante los datos anteriormente analizados, la escala obtiene propiedades métricas adecuadas, como para que la validación sea avalada por los propios resultados. Aun así, en investigaciones posteriores se deberán presentar otros trabajos que aporten evidencias de validez sobre la escala y otros hallazgos en relación a otras variables contextuales; algo que resultará imprescindible para poder validar definitivamente la escala propuesta; en especial, el análisis de la relación entre el nivel de desempeño en diversas materias y el VSE-Subjetivo debe orientar líneas prioritarias de investigación.

Asimismo, este tipo de trabajos puede ofrecer herramientas de indagación y de mejora que permitan facilitar instrumentos de evaluación con relación a la percepción que tiene el alumnado sobre la importancia de la educación. Por ello, también será necesario realizar estudios que permitan adaptar la escala hacia otras audiencias implicadas en el proceso de enseñanza-aprendizaje, otras etapas educativas e incluso a otros contextos y países.

Finalmente, en la tabla IV se adjunta la versión final del instrumento de evaluación, la escala SECS-EVALNEC VSE-Estudiantes que se estructura de manera unidimensional conformada por 20 reactivos, con una fiabilidad global de .73.

Tabla IV. Escala validada a partir de los resultados de propiedades métricas TCT y TRI

Escala EVAL/SECS VSE-Subjetivo-Estudiantes-Secundaria ⁶
1 ¿Tus profesores/as creen que vas a suspender?
2 ¿Tus padres quieren que trabajes porque estudiar no sirve para tener más dinero?
3 ¿Tus padres creen que es más importante estudiar que ganar dinero?
4 ¿Tus padres creen que vas a aprobar?
5 ¿Es importante para ti ir bien en tus estudios?
6 ¿En tu centro escolar te ayudan a tener confianza en ti mismo/a?
7 ¿En tu centro escolar te enseñan cosas útiles para un futuro trabajo?
8 ¿En tu centro escolar te ayudan a encontrar amigos/as?
9 ¿En tu centro escolar te ayudan a aprender a vivir con los/as demás?
10 ¿Lo que estás aprendiendo en tu centro te servirá cuando decidas buscar trabajo?
11 ¿Estudiar te ayuda a aprobar?
12 ¿Tienes compañeros/as que aprueban sin merecérselo?
13 ¿Piensas que las personas que estudian triunfan más?
14 ¿Las personas que triunfan en la vida no han estudiado?
15 ¿Tus padres quieren que trabajes cuanto antes porque hace falta dinero en casa?
16 ¿Cuanta más gente haya estudiado en tu ciudad mejor para todos/as?
17 ¿Ser profesor/a es una profesión muy importante?
18 ¿Tus amigos y amigas no quieren estudiar porque piensan que no sirve para nada?
19 ¿Los/as políticos/as de tu país hacen muchas cosas para que tu educación sea mejor?
20 ¿Las personas famosas normalmente no han estudiado nada?

Por todo ello, el VSE-Subjetivo podría significar una herramienta de cambio educativo que permitiera mejorar aquellas situaciones injustas que se presentan de manera subjetiva, y que se han ido constatando de manera objetiva. En este caso, hay que señalar la estigmatización consolidada que supone la vinculación, ya desde el informe Coleman et al. (1966), entre el nivel socioeconómico y cultural de las familias y el rendimiento escolar esperado en función de factores contextuales (Carabaña, 2016). Algo que potencia seguir investigando de manera que se puedan encontrar elementos que interpreten la varianza no explicada, por ejemplo en relación con la evaluación de sistemas educativos, ya que las expectativas de logro del estudiante sobre el nivel de este índice familiar no es tan directo y va perdiendo consistencia desde hace años (OECD, 2010). Por otro lado, habría que seguir investigando y realizando estudios diferenciales sobre los indicadores de evaluación con los que se trabaja para medir estos índices, que presentan tanta repercusión social y educativa (González-Such, Sancho-Álvarez y Sánchez-Delgado, 2016).

En este sentido, puede ser interesante explorar diferentes enfoques que intentan aproximar esta realidad en función de los datos obtenidos y en relación con otras variables estudiadas. Por ejemplo, resulta relevante destacar la teoría de la expectativa del valor (*expectancy-value-theory*) desarrollada por Eccles et al. (1983) sobre los determinantes de la motivación y las percepciones que tienen los estudiantes, así como sus expectativas de éxito. En este caso, un aspecto muy importante es la dimensión que plantean sobre el valor de la tarea escolar, la que desglosan en valor intrínseco, valor de utilidad personal y valor de utilidad social, algo que sin duda se puede relacionar también con la percepción del VSE-Subjetivo del alumnado.

Sobre estos supuestos nacen otros modelos teóricos sobre la medición de factores que influyen en la decisión de ser docente, como el propuesto por Watt y Richardson (2007). Asimismo, en esta línea se adaptan y validan instrumentos de evaluación para conocer los factores que influyen en la elección de los

⁶ Se recomienda su aplicación con una escala de valoración del 1 al 4. Si se implementa el instrumento de evaluación, se ruega comunicarlo a los autores de manera informativa, ya que su uso es libre si se cita adecuadamente su referencia.

estudios de educación, como el Factors Influencing Teaching (FIT-Choice) en la versión española de Gratacós y López-Jurado (2016). Todo ello nos permite seguir avanzando y enriqueciendo la investigación, así como plantear nuevas líneas de indagación o prospectivas de estudio.

En general, el VSE-Subjetivo integrado en el modelo de evaluación para la Cohesión Social desde la Educación (Jornet, 2012) se relaciona directamente con el impacto que el proceso educativo puede llegar a tener sobre una sociedad en cuestión (Jornet et al., 2015). Por tanto, desde este vector, podemos encontrar indicadores que nos permitan, a nivel objetivo, conocer los posibles orígenes de esa influencia sobre la visión de la educación a nivel subjetivo. Así, puede que las consecuencias de los niveles en estos indicadores complejos puedan estar condicionando la importancia que se le da a la educación en una sociedad, evidentemente dependiendo de cada contexto.

Referencias

- Boone, W., Staver, J. y Yale, M. (2014). *Rasch analysis in the human sciences*. Nueva York: Springer.
- Carabaña, J. (2016). El informe Coleman, 50 años después. *Revista de Sociología de la Educación*, 9(1), 9-21.
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, F., Mood, A. M., Weinfeld, F. D. y York, R. L. (1966). *Equality of educational opportunity*. Washington, DC: Government Printing Office.
- De la Orden, A. (2007). *Evaluación de la calidad de la educación. Un modelo sistémico como base para la construcción de un sistema de indicadores*. México: Instituto Nacional de Evaluación Educativa.
- Delgado, C. (2014). *Viajando a Ítaca por mares cuantitativos*. España: Amaru.
- Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M. y Meece, J. L. (1983). Expectancies, values, and academic behaviors. En J. T. Spence (Ed.), *Achievement and achievement motives. Psychological and sociological approaches* (pp. 75-146). San Francisco, CA: Freeman.
- González-Barbera, C., Caso-Niebla, J., Díaz-López, K. y López-Ortega, M. (2012). Rendimiento académico y factores asociados. Aportaciones de algunas evaluaciones a gran escala. *Bordón*, 64(2), 51-68.
- González-Montesinos, M. J. (2008). *El análisis de reactivos con el Modelo Rasch*. México: Universidad de Sonora/ Instituto Nacional de Evaluación Educativa.
- González-Montesinos, M. J. y Backhoff, E. (2010) Validación de un cuestionario de contexto para evaluar sistemas educativos con modelos de ecuaciones estructurales. *RELIEVE*, 14(2), 1-17. Recuperado de http://www.uv.es/RELIEVE/v16n2/RELIEVEv16n2_1.htm
- González-Such, J., Sancho-Álvarez, C. y Sánchez-Delgado, P. (2016). Background questionnaires of PISA: a study of the assessment indicators. *RELIEVE*, 22(1), M7. doi:10.7203/relieve.22.1.8274
- Gratacós, G. y López-Jurado, M. (2016, abril-junio). Validación de la versión en español de la escala de los factores que influyen en la elección de los estudios de educación (FIT-choice). *Revista de Educación*, 372, 87-110. doi:10.4438/1988-592X-RE-2015-372-316
- Joaristi, L., Lizasoain, L. y Gamboa, E. (2012). Construcción y validación de un instrumento de medida del Nivel Socioeconómico y Cultural (NSE). *Bordón*, 64(2), 151-171.
- Jornet, J. M. (2012). Dimensiones docentes y cohesión social: reflexiones desde la evaluación. *Revista Iberoamericana de Evaluación Educativa*, 5(1), 349-362. Recuperado de http://www.rinace.net/riee/numeros/vol5-num1_e/art27.pdf

- Jornet, J. M., López-González, E. y Tourón, J. (2012). Evaluación de sistemas educativos: teoría y experiencia. *Bordón*, 64(2), 9-12.
- Jornet, J. M., Perales, M. J. y Sánchez-Delgado, P. (2011). El valor social de la educación: entre la subjetividad y la objetividad. Consideraciones teórico-metodológicas para su evaluación. *Revista Iberoamericana de Evaluación Educativa*, 4(1), 51-77. Recuperado de <http://www.rinace.net/riee/numeros/vol4-num1/art3.pdf>
- Jornet, J. M., Sánchez-Delgado, P. y Perales, M. J. (2015). *La evaluación del impacto y la relevancia de la educación en la sociedad*. Valencia, España: PUV.
- Linacre, J. M. y Wright, B. D. (2009). *WINSTEPS: Multiple-choice, rating scale, and partial credit Rasch analysis*. Chicago, IL: MESA Press.
- Martínez-Arias, M. R., Hernández-Lloreda, M. J. y Hernández-Lloreda, M. V. (2006). *Psicometría*. Madrid: Alianza Editorial.
- Ministerio de Educación, Cultura y Deporte. (2013). *Panorama de la Educación. Indicadores de la OCDE 2013. Informe español*. Madrid: Autor.
- Moliner, O., Traver, J. A., Ruíz, M. P. y Segarra, T. (2016). Estrategias que indican en los procesos de democratización de la escuela. Una aproximación teórica. *Revista Electrónica de Investigación Educativa*, 18(2), 116-129. Recuperado de <http://redie.uabc.mx/redie/article/view/1110>
- OCDE. (2010). *Regards sur l'éducation 2010* [Un vistazo a la educación 2010]. París : Autor.
- OCDE. (2013). *Education at a Glance 2013: OECD Indicators*. París : Autor.
- Sancho-Álvarez, C., Jornet, J. M. y González-Such, J. (2016). El constructo Valor Social Subjetivo de la Educación: validación cruzada entre profesorado de escuela y universidad. *Revista de Investigación Educativa*, 34(2), 329-350.
- Sancho-Álvarez, C., Jornet, J. M. y González-Such, J. (2017). Design of scales to assess the Subjective Social Value of Education for students in primary and secondary school. *Procedia – Social and Behavioral Sciences*, 237C, 527-534. doi:10.1016/j.sbspro.2017.02.102
- Valcárcel, A. (2013). *Igualdad, educación y mundo compartido*. Conferencia impartida en las XVII Jornadas de Información para orientadores. Universitat de València, España.
- Watt, H. M. G. y Richardson, P. W. (2007). Motivational factors influencing teaching as a career choice: development and validation of the FIT-choice scale. *The Journal of Experimental Education*, 75(3), 167-202.