



¿Qué relación existe entre las ideas de los maestros en formación inicial sobre indagación y evaluación en ciencias?

María José Sáez Bondía^{1a}, Ángel Luis Cortés Gracia^{1b}, María José Gil Quílez^{1b},
Begoña Martínez Peña^{1b} y Santos Orejudo Hernández^{2b}

¹Departamento de Didáctica de las Ciencias Experimentales. ² Departamento de psicología y Sociología. ^a Facultad de Ciencias Humanas y de la Educación. ^b Facultad de Educación. Universidad de Zaragoza. España

[Recibido el 29 de octubre de 2018, aceptado el 30 de abril de 2019]

Desde la investigación educativa son muchos los trabajos que vinculan, desde una perspectiva cualitativa, los modelos docentes predominantes con el modo en el que se evalúan los aprendizajes en las clases de ciencias. Este trabajo, desde un enfoque cuantitativo, analiza la relación existente entre los perfiles de los maestros ante la enseñanza de las ciencias por indagación y sus creencias sobre cómo se debería evaluar en las clases de ciencias. Para ello, se han diseñado y validado dos escalas (BAIST y BASLA) a partir de las cuales se analiza esta relación. Dichas escalas fueron administradas a un total de 255 estudiantes del grado de maestro en educación primaria. Tras la realización de un ANOVA factorial entre los perfiles obtenidos y las respuestas dadas en el cuestionario sobre evaluación se observa una clara relación entre los perfiles de los maestros en formación inicial ante la indagación y el modo en el que evaluarían los aprendizajes de sus estudiantes.

Palabras clave: evaluación; creencias; formación inicial del profesorado; indagación; cuestionarios

What relationship does it exist in pre-service teachers' ideas between inquiry and assessment in science education?

From educational research there are many works that link, from a qualitative perspective, the predominant teaching models with the way in which learning is evaluated in science classrooms. This work analyzes, from a quantitative approach, the relationship between teachers' profiles about teaching of science by inquiry and their beliefs about how science learning should be assessed. To do this, two scales (BAIST and BASLA) have been designed and validated. These scales were applied to a total amount of 255 students of the university grade of teacher in primary education. After carrying out a factorial ANOVA between the profiles obtained and the answers given in the questionnaire on evaluation, a clear relation is observed between the pre-service teachers' profiles about inquiry and the way in which they would assess the learning of their students.

Keywords: assessment; beliefs; pre-service teacher training; inquiry; questionnaires

Para citar el artículo. Sáez Bondía, M.J., Cortés Gracia, A.L., Gil Quílez, M.J., Martínez Peña, B. y Orejudo Hernández, S. (2019). ¿Qué relación existe entre las ideas de los maestros en formación inicial sobre indagación y evaluación en ciencias?. *Ápice. Revista de Educación Científica*, 3(2), 93-108. DOI: <https://doi.org/10.17979/arec.2019.3.2.4631>

Contacto: msaezbo@unizar.es, acortes@unizar.es, quilez@unizar.es, bpena@unizar.es, sorejudo@unizar.es

Introducción

Las creencias del profesorado en formación sobre la enseñanza y aprendizaje de las ciencias experimentales tienen gran repercusión en el desarrollo de su futura práctica docente, incluida su predisposición a posibles cambios pedagógicos (Porlán y Martín del Pozo, 2004; Markic, Eilks y Valanides, 2008; Cortés et al., 2012; Garritz, 2014; Sáez Bondía et al., 2017; Hamilton, 2018). En las últimas décadas podemos encontrar diversos trabajos que analizan esas creencias, sobre temas generales o específicos, tanto de maestros en ejercicio como de maestros en formación inicial, sobre la naturaleza de la ciencia o sobre cómo enseñar ciencias y sus implicaciones en el proceso de enseñanza y aprendizaje (Pajares, 1992; Hashweh, 1996; Bryan 2003, Sampson y Benton, 2006; Campbell, Abd-Hamid y Chapman, 2010, Martín, Prieto y Jiménez, 2013; García Carmona y Acevedo Díaz, 2016; Yoon y Kim, 2016, entre otros).

En los últimos veinte años, buena parte de las recomendaciones sobre cómo abordar la enseñanza y el aprendizaje de las ciencias hacen referencia a las actividades de indagación o investigación escolar, de modo que la formación del profesorado no debería ser ajena a ellas (National Research Council, 2000; Windschitl, 2003; Flick and Lederman, 2006; Cañal, Travé y Pozuelos, 2011; Forbes, 2011). Sin embargo, diversos autores indican diferentes tipos de dificultades con las que se encuentra el profesorado para poner en marcha estas estrategias (Anderson, 2002; Kirschner, Sweller y Clark, 2006; Gil Quílez et al., 2008; Couso, 2014; Romero Ariza, 2017). Por ello, las creencias de los profesores en formación y en ejercicio sobre la indagación también jugarán un papel fundamental en la posible toma de decisiones pedagógicas de los mismos acerca del tema (Wallace y Kang, 2004; Abril et al., 2014; van Uum, Verhoeff y Peeters, 2016; van Aalderen-Smeets et al., 2017, entre otros).

De la misma manera que las creencias de los maestros sobre las ciencias y su enseñanza juegan un papel fundamental en su práctica docente, las creencias sobre cómo deben ser evaluados los aprendizajes del alumnado también tienen relación directa con el modo en que los docentes conciben el proceso de enseñanza y aprendizaje (Ibarra y Rodríguez, 2010). De hecho, como indican Rodríguez Barreiro, Gutiérrez Muzquiz y Molledo (1992, p. 254), “cualquier intento de mejora de los procesos de enseñanza-aprendizaje está condenado al fracaso -o falta de operatividad- si no va acompañado, paralelamente, de un perfeccionamiento en los modelos y técnicas de evaluación”, ya que existe una íntima relación entre el modelo de enseñanza y la evaluación (Atkin, Black y Coffey, 2001).

A pesar de lo anterior, y aunque existen desde hace tiempo trabajos sobre la repercusión de las creencias sobre evaluación en el desarrollo profesional de los docentes (McMillan, Myran y Workman, 2002; Brown, 2004; Barnes, Fives y Dacey, 2017), son escasos los estudios que se centran en las creencias sobre la evaluación de los aprendizajes en el contexto de las materias científicas (Alonso, Gil y Martínez-Torregrosa, 1996; Bryan, 2012; Pontes Pedrajas, Poyatos López y Oliva Martínez, 2016).

De este modo, el objetivo principal de este trabajo consiste en determinar, desde un punto de vista cuantitativo, la existencia de una relación en las creencias sobre evaluación e indagación en la formación inicial de maestros. Para ello, se parte del diseño y validación de dos escalas: una relacionada con las creencias sobre cómo enseñar ciencias (Sáez Bondía, Cortés Gracia, Gil Quílez, Martínez Peña y Orejudo Hernández, 2017) y otra sobre evaluación. A partir del análisis descriptivo de las mismas y del establecimiento de perfiles de indagación a partir de la primera escala, se analiza esta posible relación.

Marco teórico

Las creencias sobre evaluación suponen teorías implícitas de los maestros sobre la enseñanza y el aprendizaje, sobre la cultura escolar predominante, sobre la organización en el

aula y sobre la estructura cognitiva de maestros y alumnos (Pozo, 2000, Bryan, 2012). Por tanto, tienen una relación con los modelos docentes predominantes, pudiendo encontrar, de forma simplificada, dos tipos de maestros: 1) los que creen que la evaluación es un instrumento de control que proporciona información objetiva sobre las respuestas correctas (Stiggins, 2006) y que, de este modo, pueden identificar los “errores” de los estudiantes y reconducirlos con la realización de más ejercicios (Sanmartí, 2007) y 2) los que piensan que la evaluación permite detectar las causas de las dificultades de los estudiantes y se interesan más por los procesos en el desarrollo de tareas por parte del alumnado que por los resultados finales. El primer grupo correspondería a una visión tradicional que considera las producciones finales (Porlán, 1993). El segundo grupo poseería una visión más cognitivista de la evaluación (Sanmartí, 2007).

Estas dos visiones de la evaluación se relacionan con los modelos didácticos (Gil Pérez, 1983; Porlán y Rivero, 1998; Jiménez Aleixandre, 2000), es decir, con las visiones existentes sobre los procesos de enseñanza y aprendizaje, la innovación educativa e incluso con una posible resistencia ante cualquier tipo de cambio (Campanario y Moya, 1999). Dependiendo de los autores podemos encontrar diferentes tipificaciones de los modelos que, de alguna manera, se integrarían en las categorías propuestas por autores como Gil Pérez (1983) o Jiménez Aleixandre (2000): a) modelo de transmisión-recepción (o tradicional), b) modelo de descubrimiento (autónomo o espontaneísta) y c) modelo constructivista de aprendizaje (de indagación dirigida o de investigación en la escuela).

El modelo tradicional de enseñanza se relaciona con evaluaciones lo más reproductivas posibles y se centra en listas de conceptos, tendiendo a evaluaciones muy específicas o puntuales del conocimiento. En cuanto a la evaluación asociada al modelo de aprendizaje por descubrimiento, aunque centra su visión en los procesos, también tiene en consideración los productos finales (Pozo y Gómez Crespo, 1998). No obstante, en el caso de la formación inicial de maestros, parece que los enfoques relacionados con este modelo docente están poco vigentes (Lucha, Bravo Torija, Forcadell y Ferrer, 2016). Por tanto, en este caso se podrían considerar las variaciones vinculadas a un modelo de indagación dirigida (Windchitl, 2003; NRC, 2000), atendiendo al nivel de autonomía en el aprendizaje de los estudiantes. Desde la perspectiva de la evaluación de los aprendizajes, este enfoque se enmarcaría dentro de un modelo más cognitivista.

Si nos centramos en la formación inicial del profesorado, parece que las preocupaciones en cuanto a la evaluación de los aprendizajes son mucho menores que en el caso del profesorado en activo cuando se les pide que diseñen propuestas (Gil Quílez, Martínez Peña y Cordero, 2017) o cuando se les pregunta por sus necesidades formativas (Cortés et al., 2012).

No obstante, independientemente de la importancia que los maestros en formación inicial atribuyan a la evaluación en determinadas situaciones, suponemos que debería existir algún tipo de relación entre la visión que posee el profesorado sobre la enseñanza y aprendizaje y la función que se otorga a la evaluación y que esta relación debería ser coherente desde el punto de vista de la futura práctica docente.

Método

Participantes y contexto

La muestra utilizada para el desarrollo del estudio consistió en 255 estudiantes del tercer curso del Grado de Maestro en Educación Primaria que cursaban la asignatura de Didáctica del Medio Biológico y Geológico en la Universidad de Zaragoza. Los estudiantes

procedían de los tres campus de la universidad de estudio: 56.1% (n=143) de Zaragoza, 12.2% (n=31) de Huesca y 31.8% (n=81) de Teruel.

Estos estudiantes habían cursado previamente una asignatura relacionada con la didáctica de las ciencias, aunque los aspectos relativos a la enseñanza por indagación eran introducidos durante la citada asignatura de tercer curso.

La edad media de los estudiantes participantes en el estudio era de 22.3 años, siendo un 65.9% (n=168) mujeres y un 34.1% (n=87) hombres. Para el presente estudio se tuvieron en consideración las asignaturas de ciencias cursadas durante la etapa de bachillerato, observándose que el 77.3% de los estudiantes participantes habían cursado tres o menos asignaturas de ciencias de las siete posibles a cursar en esa etapa.

Procedimiento de diseño de los instrumentos

Para conocer la relación existente entre las ideas sobre indagación y evaluación de los maestros en formación inicial se diseñaron dos cuestionarios. Uno relacionado con las creencias a la hora de enseñar ciencias denominado como escala BAIST (Beliefs About Inquiry Science Teaching o Creencias sobre la enseñanza de las ciencias por indagación) y otro relacionado con las creencias sobre evaluación denominado como escala BASLA (Beliefs About Science Learning Assessment o Creencias sobre la evaluación de los aprendizajes en ciencias). A continuación se describe el proceso de diseño, validación y obtención de los cuestionarios y las técnicas de análisis de datos empleadas para conocer la relación existente entre ambos tipos de creencias.

Construcción de los ítems y evaluación

La elaboración de los ítems para ambos cuestionarios parte de un trabajo anterior en el que los maestros en formación inicial declaraban sus expectativas y necesidades formativas (Cortés et al., 2012). En concreto, a través de un cuestionario abierto se exploraban las ideas de los estudiantes sobre cinco dimensiones relacionadas con la enseñanza de las ciencias, centrándonos en las tres primeras para este trabajo: 1) los conocimientos que pensaban que eran necesarios para enseñar ciencias; 2) los aspectos que consideraban fundamentales para llevar a cabo actividades de indagación en las aulas y 3) las evidencias de la existencia de aprendizajes en los estudiantes tras realizar actividades de indagación (aspectos evaluativos).

Las respuestas de los estudiantes fueron categorizadas para cada una de las cuestiones planteadas. Tras una revisión detallada del contenido de las mismas, se fueron redactando frases representativas que constituirían posibles ítems de las escalas finalmente diseñadas. En concreto, para la escala BAIST se tuvieron en consideración las respuestas relativas a las cuestiones 1 y 2, mientras que para la escala BASLA se consideraron las respuestas dadas ante la cuestión 3.

El cuestionario BAIST piloto estaba constituido por 44 ítems de escala Likert 1-4, siendo 1 "totalmente en desacuerdo", 2 "desacuerdo", 3 "acuerdo" y 4 "totalmente de acuerdo" (valoradas con puntuaciones del 1 al 4). Las frases iniciales fueron refinadas y evaluadas por un grupo de expertos en cuanto a su comprensión y claridad. Asimismo, fueron agrupadas en tres dimensiones iniciales vinculándolas al constructo teórico relativo a la enseñanza de las ciencias en contextos de indagación. El primer factor incluía aquellas frases que hacían referencia a los contenidos de ciencias necesarios para enseñar ciencias por indagación; el segundo a los aspectos metodológicos (diseño y aplicación de propuestas en contextos de indagación) y el tercero a las posibles limitaciones presentes a la hora de llevar a cabo actividades de indagación en las aulas de primaria.

Para el diseño del cuestionario BASLA piloto se procedió de un modo similar. En este caso, el cuestionario inicial estuvo constituido por 27 ítems de escala Likert 1-4. Dichos ítems, derivados de la creación de frases representativas refinadas a partir de las respuestas dadas ante la cuestión 3 del estudio previo, fueron agrupados en dos dimensiones asociadas a la evaluación de los aprendizajes en las clases de ciencias. La primera categoría incluía frases relativas a una visión tradicional de la evaluación, la segunda consideraba una visión cognitivista o constructivista de la evaluación.

Finalmente, con objeto de validación, ambas escalas piloto fueron administradas a los estudiantes participantes del estudio, previa autorización de los mismos, el primer día de clase. Estas escalas fueron administradas de forma separada (primero cumplieron la escala BAIST y a continuación la escala BASLA). El tiempo de cumplimentación de ambas escalas fue aproximadamente de una hora. Asimismo, en los cuestionarios se incluían preguntas relativas a los datos sociodemográficos de los participantes, tales como sexo, edad y estudios previos y se demandaba que incluyesen un código que les identificase de forma anónima (código universitario) que serviría para comparar resultados tanto de ambas escalas como de posibles cambios en sus respuestas en estudios longitudinales futuros.

Validación de las escalas

El primer paso para la validación de las escalas fue la realización de un análisis descriptivo seguido del estudio de la consistencia interna. A continuación, se realizó un análisis factorial exploratorio para cada una de las escalas con el uso del programa informático de análisis de datos M-PLUS (Muthén y Muthén, 1998-2010). Este programa permite la estimación del número adecuado de factores a partir de los datos y facilita tanto la realización de análisis factoriales exploratorios (AFE) como confirmatorios (AFC). Los procedimientos estadísticos empleados para la realización de ambos tipos de análisis se basan en el desarrollo de modelos exploratorios de ecuaciones estructurales (método ESEM, inglés Exploratory Structural Equation Modeling). Este método permite estimar tanto el número de factores más adecuado en una estructura como los pesos factoriales y ofrece alternativas a los modelos de extracción de factores clásicos, tales como el de Kaiser (Lloret-Segura, Ferreres-Traver, Hernández-Baeza y Tomás-Marco, 2014). Asimismo, permite usar métodos de estimación robustos como el de máxima verosimilitud robusta (R-MLE).

Para analizar los valores de ajuste de los distintos modelos se tomaron los indicadores aportados por este programa, tales como el valor de Chi-cuadrado, CFI, TLI o root mean square (RMSEA), adoptando criterios de valoración de los mismos como los incluidos en Brown (2006).

El AFE se llevó a cabo usando el método de máxima verosimilitud (MLE maximum likelihood estimation, en inglés) que permitió la obtención de un modelo matemático para cada una de las escalas. Este modelo inicial estaba conformado por cuatro factores en ambas escalas.

A continuación, atendiendo a la clasificación teórica realizada en la fase de construcción de los ítems, se llevó a cabo un AFC utilizando el método de máxima verosimilitud robusto (R-MLE). Se obtuvo un modelo teórico de cuatro factores para la escala BAIST y uno de tres factores para la escala BASLA.

El número de ítems de cada uno de los modelos creados se redujo siguiendo los siguientes criterios: 1) los pesos factoriales ítem-factor tenían que ser superiores a 0.4 y 2) aquellos ítems en los que las puntuaciones de sus percentiles 25, 50 y 70 eran las mismas serían eliminados ya que este tipo de ítems discriminaban poco entre participantes (Brown, 2006).

En ambos procesos de validación de las escalas nos encontrábamos con la problemática

de que el modelo matemático (realizado con el AFE) daba mejores valores estadísticos que el teórico (realizado con el AFC), pero no lograba ser agrupado atendiendo a un constructo teórico coherente. Por este motivo se optó por realizar un modelo mixto que tratase de solucionar ambos problemas.

Los ítems coincidentes para cada factor en el modelo matemático y teórico se tuvieron en consideración para elaborar el modelo mixto tanto para la escala BAIST como para la escala BASLA. Una vez creados los dos modelos mixtos, se llevó a cabo un nuevo AFC a través de un MLE y se utilizaron los mismos criterios de exclusión de ítems que con el resto de modelos obtenidos. La escala mixta fue la empleada como escala definitiva para los dos cuestionarios, ya que solucionaba los problemas descritos anteriormente.

Cuestionarios finales

La primera versión de la escala BAIST se redujo de 44 a 20 ítems. Este modelo final estaba compuesto por los ítems coincidentes de los dos modelos anteriores y aquellos ítems que tenían mejores ajustes. Los datos de ajuste del modelo mixto fueron mejores que los obtenidos en los modelos previos (RMSEA=0.048; CFI=0.929; TLI=0.914) y eran más acordes al constructo teórico. Así, el modelo final quedó agrupado en cuatro factores: 1) saber conceptual y metodológico; 2) los aspectos a considerar en la planificación de actividades de ciencias; 3) las competencias del profesor en la aplicación de las actividades de ciencias y 4) aspectos específicos a considerar en la enseñanza de las ciencias.

En el caso de la escala BASLA, los datos del ajuste obtenidos (RMSEA=0.048; CFI=0.929; TLI=0.914) eran mejores que los obtenidos en el modelo teórico, pero peores que el matemático. No obstante, los valores de ajuste estaban dentro de los establecidos por Brown (2006) y eran más concordantes con el constructo teórico relativo a la evaluación de los aprendizajes en ciencias. En este caso la escala final pasó de 27 a 16 ítems agrupados en tres factores: 1) valora exclusivamente las producciones finales del alumnado; 2) valora las tareas y producciones que implican un alto nivel de autoaprendizaje por parte del alumnado y 3) valora las tareas y producciones que implican un nivel intermedio de autoaprendizaje por parte del alumnado.

Las tablas 1 y 2 muestran los cuestionarios finales con sus respectivos coeficientes para cada factor e ítem.

Análisis de la relación entre las ideas sobre indagación y evaluación a partir de las escalas BAIST y BASLA definitivas

A partir de los cuestionarios validados y refinados se llevó a cabo un análisis descriptivo de las respuestas dadas por los estudiantes y para ello se empleó el programa de análisis de datos SPSS 19. Los valores establecidos se promediaron con el uso de una horquilla 1 a 4 atendiendo a los valores de la escala Likert planteada en el cuestionario.

En el caso de la escala BAIST se llevó a cabo un análisis de conglomerados de k medias (análisis cluster). Éste análisis nos permitió agrupar a los estudiantes atendiendo a un perfil determinado ante la enseñanza de las ciencias por indagación, considerando la diferencia de medias para cada uno de los factores obtenidos en el cuestionario final fuese pequeña en cada conglomerado (tomando una distancia euclídea). En el caso de la escala BASLA no aparecían agrupaciones diferenciadas a la hora de llevar a cabo el análisis de conglomerados, así que no se consideró.

Los perfiles obtenidos en el caso de la escala BAIST fueron tres (Sáez Bondía et al., 2017): 1) el perfil A (n=85), denominado como “más indagador” con altas puntuaciones en los tres primeros factores de la escala. Estudiantes que consideran necesarios tanto conoci-

Tabla 1. Cuestionario final sobre creencias ante la enseñanza por indagación: escala BAIST. Adaptado de Sáez Bondía et al. (2017)

Ítem	Coeeficientes	Alfa ¹
<i>Factor 1: Saber conceptual y metodológico</i>		0.725
Dominar técnicas para realizar actividades en el medio natural y en el laboratorio	0.644	
Dominar técnicas para realizar con sus alumnos actividades en el medio natural, en el aula y/o en el laboratorio	0.620	
Elaborar propuestas de actuación para la innovación en las aulas, acorde con los cambios en el ambiente social	0.563	
Cómo manejar aparatos de uso común en clases de ciencias naturales	0.512	
Conocer los problemas del mundo actual que inciden sobre aspectos del medio natural	0.664	
<i>Factor 2: Aspectos a considerar en la planificación de actividades de ciencias</i>		0.708
Adaptar los contenidos de ciencias al nivel de comprensión del alumnado de primaria	0.577	
Conocer las capacidades propias de las distintas etapas de desarrollo en las que se encuentran sus alumnos	0.712	
Diagnosticar cuáles son las dificultades que cada alumno tiene cuando se enfrenta a una actividad	0.618	
Captar y mantener la atención de los alumnos durante la clase	0.564	
<i>Factor 3: Competencias del profesor en la aplicación de actividades de ciencias</i>		0.717
Ser capaz de prever cuáles serían las dificultades procedimentales que el alumnado tendría a lo largo de una actividad.	0.445	
Ser capaces de plantear preguntas que inciten a los alumnos a elaborar respuestas que no figuran en los libros de texto	0.504	
Diseñar actividades para el aula que ayuden a los alumnos a construir y a comprender la teoría sobre el medio natural	0.509	
Motivar a los alumnos para que formulen hipótesis que puedan probar en clase	0.572	
Ser capaz de enseñar habilidades de comunicación para trabajar en las clases de ciencias	0.577	
Diseñar actividades que ayuden al alumno a comprender y a razonar sobre las informaciones que recibe en su vida cotidiana, desde la ciencia	0.557	
Organizar el trabajo en equipo, de manera que todos los alumnos de la clase participen en la construcción de conocimientos de ciencias	0.458	
<i>Factor 4: Aspectos específicos a considerar en la enseñanza de las ciencias</i>		0.601
Que las investigaciones en el aula no son posibles sin instrumental científico	0.537	
Tener en cuenta que los fenómenos naturales no se pueden reproducir en el aula de Primaria	0.512	
Los contenidos del currículum de primaria permiten la realización de actividades de investigación en el aula	0.439	

Tabla 1. Cuestionario final sobre creencias ante la enseñanza por indagación: escala BAIST. Adaptado de Sáez Bondía et al. (2017). *Continuación*

Ítem	Coefficientes	Alfa ¹
Un impedimento para investigar en la escuela es que la mayor parte del instrumental científico requerido no es apto para que lo manipulen los niños	0.585	

¹Alfa de Cronbach para cada factor

Tabla 2. Cuestionario final sobre creencias sobre evaluación: escala BASLA

Ítem	Coefficientes	Alfa ¹ de Cronbach
Factor 1: Valora exclusivamente las producciones finales del alumnado		0.686
Reproduce literalmente los conceptos enseñados cuando se lo demandan	0.672	
Responde correctamente en sucesivas pruebas escritas	0.744	
Por encima de lo que aparenta saber en clase, es imprescindible una prueba escrita para poder apreciar su nivel de conocimientos	0.504	
Reproduce lo esencial de lo trabajado en los cuadernos de clase, o portafolios	0.512	
Factor 2: Valora tareas y producciones que implican un alto nivel de autoaprendizaje por parte del alumnado		0.734
Trabaja en equipo, pide explicaciones a sus compañeros acerca de decisiones que no coinciden con las suyas	0.458	
Formula preguntas que pudieran ser investigadas científicamente	0.581	
Responde a las cuestiones planteadas empleando expresiones propias	0.518	
Intenta resolver los problemas planteados en clase, utiliza conceptos científicos relevantes	0.498	
Selecciona preguntas que pueden ser contestadas/evaluadas mediante investigación científica	0.725	
Contrasta los resultados obtenidos con los conocimientos previos a la hora de establecer conclusiones	0.534	
Factor 3: Valora tareas y producciones que implican un nivel intermedio de autoaprendizaje por parte del alumnado		0.652
Las preguntas que plantea no sólo se refieren a datos (nombres, fechas, tamaños, etc.)	0.484	
Pide aclaraciones al profesor cuando no comprende lo que éste expone	0.582	
Propone soluciones que se pueden comprobar	0.655	
Expresa con sus propias palabras los problemas	0.576	
Selecciona los datos que necesita para comprobar su solución al problema	0.744	
Busca información relevante para poder resolver los problemas planteados	0.772	

¹Alfa de Cronbach para cada factor

mientos y competencias por parte del maestro para llevar a cabo actividades de ciencias y dan importancia a una planificación previa de estas actividades. Es el grupo que menos limitaciones percibe a la hora de aplicar actividades en las aulas de primaria (puntuaciones bajas en el factor 4); 2) el perfil B (n=87), con un perfil "intermedio", presenta valores similares al grupo A en los tres primeros factores y valores más altos en el factor 4 y; 3) el perfil C (n=71) que presenta valores más bajos en los tres primeros factores y un valor intermedio entre los perfiles A y B para el factor 4. Este último grupo, fue denominado como perfil "pesimista" ante la indagación.

A continuación se realizó una primera comparación entre la escala BAIST Y BASLA utilizando la correlación de Spearman entre los factores de cada una de las escalas. Por último, se realizó un ANOVA factorial entre cada uno de los perfiles de los estudiantes, obtenidos a partir del análisis de conglomerados de la escala BAIST y los factores de la escala BASLA. Este análisis estudia las relaciones entre factores agrupados en perfiles de la escala BAIST y la escala BASLA, dando información sobre las relaciones entre perfiles ante la enseñanza de las ciencias por indagación y sus creencias sobre evaluación, aspecto que no se discrimina en el estudio de las correlaciones entre factores.

Resultados

Resultados descriptivos de los cuestionarios BAIST Y BASLA

Una vez validadas las escalas BAIST y BASLA se llevó a cabo un estudio descriptivo de las respuestas dadas por los estudiantes utilizando la media de las respuestas para cada uno de los factores obtenidos en cada cuestionario.

Globalmente, los maestros en formación inicial consideran muy importante para enseñar ciencias: 1) tener conocimientos tanto sobre la materia objeto de estudio como sobre metodologías (factor 1 de la escala BAIST: n=251; \bar{x} =3.51, Desviación estándar (D.E.)=0.38), 2) planificar las actividades de ciencias (factor 2 de la escala BAIST: n=253; \bar{x} =3.53, D.E.=0.39) y 3) tener competencias a la hora de aplicar determinadas actividades en la clase de ciencias (factor 3 de la escala BAIST: n=253; \bar{x} =3.54, D.E.=0.33). Sin embargo, las puntuaciones relativas a los aspectos específicos, principalmente relacionados con "limitaciones" a la hora de llevar a cabo actividades de ciencias, se sitúan en un nivel intermedio, existiendo una mayor variabilidad en sus respuestas (factor 4 de la escala BAIST: n=247; \bar{x} =2.22, D.E.=0.57).

En relación a los resultados descriptivos de la escala BASLA, la puntuación atribuida a cada uno de los factores es diferente. La evaluación de los aprendizajes en la que se consideran tanto los procesos como los productos vinculados a un nivel intermedio de autonomía (factor 3) es la mejor valorada (n=255; \bar{x} =3.43, D.S.=0.35), seguida de aquella en la que se consideran niveles mayores de autoaprendizaje (n=254; \bar{x} =3.30, D.S.=0.35). Por su parte, el factor 1, en el que solo se consideran las producciones finales de los estudiantes, obtiene puntuaciones mucho más bajas y mucha más variabilidad en sus respuestas (n=252; \bar{x} =2.29, D.E.=0.57).

Relación entre creencias ante la enseñanza de las ciencias por indagación y evaluación

Correlación entre los factores de ambas escalas

La tabla 3 muestra las correlaciones entre los factores obtenidos en ambas escalas. Existe una correlación notable entre las competencias del maestro (factor 3 de la escala BAIST) y la evaluación considerando niveles medios y altos de autonomía (factores 3 y 2 de la

escala BASLA). Asimismo, se observa una asociación entre los factores 1 y 2 de la escala BAIST y los factores 2 y 3 de la escala BASLA. Además, existe cierta correlación entre las limitaciones o aspectos a considerar (factor 4 de la escala BAIST) y la evaluación basada únicamente en las producciones finales de los estudiantes (factor 1 de la escala BASLA). Al mismo tiempo, este factor de la escala BASLA tiene una débil relación negativa con el factor 3 de la escala BAIST.

Tabla 3. Correlaciones de Spearman entre los factores de cada escala

		Escala BASLA		
		Factor 1	Factor 2	Factor 3
Escala BAIST	Factor 1	-.045	.098	-.125(*)
	Factor 2	.336(**)	.328(**)	.415(**)
	Factor 3	.287(**)	.333(**)	.409(**)
	Factor 4	-.045	.098	-.125(*)

*La correlación es significativa l nivel 0.05 (bilateral)

**La correlación es significativa al nivel de 0.01 (bilateral)

Comparación entre perfiles ante la indagación y creencias sobre evaluación

La tabla 4 muestra para cada perfil de estudiantes obtenidos de la escala BAIST la importancia atribuida a los factores sobre evaluación obtenidos en la escala BASLA. En los factores en los que se observan mayores diferencias, en términos del tamaño del efecto para cada perfil, se encuentran la evaluación que implica altos niveles de autonomía o factor 3 ($\eta^2=.117$) y la evaluación con niveles intermedios de autonomía o factor 2 ($\eta^2=.067$).

Tabla 4. Comparaciones de medias entre perfiles (Sáez Bondía et al., 2017) y factores sobre evaluación con ANOVA factorial

	Conglomerado	N	Media	D.t.	F	η^2
Factor 1	1 – Perfil A	84	2.17	0.62	4.204*	.034
	2 – Perfil B	86	2.42	0.56		
	3 – Perfil C	69	2.30	0.49		
Factor 2	1 – Perfil A	84	3.42	0.40	8.502*	.067
	2 – Perfil B	86	3.38	0.41		
	3 – Perfil C	69	3.16	0.42		
Factor 3	1 – Perfil A	84	3.58	0.30	15.568*	.117
	2 – Perfil B	86	3.46	0.33		
	3 – Perfil C	69	3.29	0.35		

* $p<.01$

Como se puede observar, los perfiles de los estudiantes analizados ante la enseñanza por indagación tienen una relación evidente con sus creencias relativas a cómo evaluar el aprendizaje de las ciencias. El perfil A, que representa a los estudiantes más optimistas ante la indagación, valora con puntuaciones más bajas la evaluación tradicional. Los estudiantes con un perfil B, que perciben más limitaciones que los anteriores ante la puesta en marcha de actividades de indagación, obtienen puntuaciones intermedias. El perfil C, o más pesimista ante la indagación, da mayor peso a una evaluación tradicional.

Discusión

Si bien, desde una perspectiva cualitativa, se ha hablado en diversos trabajos sobre la existencia de una relación entre los perfiles profesionales y sus ideas sobre cómo evaluar (Atkin, Black y Coffey, 2001, Myran y Workman, 2002, Rivero et al., 2017, entre otros), apenas se ha tratado de evidenciarlo desde una perspectiva cuantitativa.

Este trabajo no solo muestra el proceso de diseño y validación de dos cuestionarios para conocer las creencias del profesorado en formación ante la enseñanza de las ciencias por indagación y su evaluación, sino que además confirma que existe una relación significativa entre ambos aspectos.

El proceso de validación de una escala no es sencillo. Así, en muchas ocasiones, las agrupaciones que realizan los análisis factoriales, basadas en el tipo de respuestas dadas por los estudiantes, se alejan de los constructos o marcos teóricos empleados desde la investigación. Este hecho, aunque a priori sea negativo, nos aporta una serie de informaciones muy relevantes desde nuestro punto de vista. El proceso de validación a través de la realización de análisis factoriales toma en consideración las respuestas y las agrupa atendiendo a la similitud de varianza de cada sujeto, es decir, atendiendo a la consistencia interna de las respuestas de los mismos. De este modo, los factores establecidos considerando los valores óptimos desde el punto de vista estadístico nos permiten hacernos una idea inicial de los marcos de referencia empleados por los participantes en el estudio. Por ejemplo, en la escala BAIST a la hora de realizar el análisis factorial exploratorio, nos encontrábamos con la problemática de que los estudiantes de magisterio no discernían entre saberes conceptuales y metodológicos, ni diferenciaban entre diseñar propuestas y aplicarlas en el aula.

Así, la solución propuesta en ambos casos fue la realización de modelos mixtos que tuviesen buenos indicadores estadísticos al mismo tiempo que se ajustaban a un modelo teórico coherente desde la investigación en didáctica de las ciencias. Por tanto, este procedimiento mostrado puede ser transferible a cualquier otra validación de cuestionarios en la que el investigador se encuentre con esta dificultad.

Además, el cuestionario BAIST no solo nos ha permitido conocer las creencias del alumnado sobre lo que se necesita saber para enseñar ciencias, sino que, a partir del mismo, hemos establecido perfiles docentes ante su enseñanza por indagación atendiendo a las puntuaciones en los diferentes factores del cuestionario. En cualquier caso se observa que solo un 35 % de los alumnos encuestados se ajustan al perfil "más indagador" (Sáez Bondía et al., 2017).

En cuanto a la relación entre las creencias sobre cómo enseñar ciencias y cómo evaluar los aprendizajes de los estudiantes, se observa una asociación relativamente coherente entre los factores obtenidos en ambas escalas. Este resultado, no solo permite evidenciar de nuevo cierta relación entre estos dos elementos sino validar la coherencia interna de los dos instrumentos creados.

La realización de un ANOVA factorial entre los perfiles obtenidos en la escala BAIST y los factores relativos a evaluación permiten confirmar, desde un punto de vista estadístico, la relación entre las creencias sobre evaluación y sobre cómo enseñar ciencias atendiendo a los perfiles creados. Por ejemplo, perfiles más pesimistas ante la enseñanza de las ciencias por indagación (se centran en las limitaciones del modelo), son los que están más de acuerdo con evaluar a los estudiantes únicamente a través de sus producciones finales. Por el contrario, los más optimistas ante el modelo indagatorio están más abiertos a valorar tareas y producciones realizadas con cierto nivel de autonomía.

Conclusiones, limitaciones y perspectivas de futuro

Los instrumentos creados con el objetivo de conocer las ideas de los maestros en formación inicial en cuanto a cómo enseñar ciencias y cómo evaluarla son un punto de partida para ampliar el estudio a otras universidades. Existen muchos cuestionarios acerca de las creencias sobre cómo enseñar ciencias, pero apenas hay instrumentos en lo relativo a como evaluar, por lo que, en este caso, las escalas BASLA Y BAIST se complementan.

No obstante, existen determinadas limitaciones en los instrumentos que habría que tener en consideración para futuros estudios. Por ejemplo, el uso de una escala Likert 1-4 podría generar un efecto techo en futuros estudios longitudinales. Si estas escalas se emplean con el objetivo de ver la evolución de las creencias a lo largo de su formación e inicialmente, tal y como se muestra en los resultados descriptivos, para alguno de los factores ya puntúan por encima de tres, puede suceder que apenas se observen variaciones por este hecho. Por tanto, y siguiendo las recomendaciones de Bryan (2012) en cuanto a la realización de estudios longitudinales, sería necesario ampliar la escala Likert al menos en un rango del 1 al 6.

La relación existente entre las creencias sobre evaluación y sobre cómo enseñar ciencias concuerdan con otros estudios de tipo cualitativo (Ibarra y Rodríguez, 2010). Este hecho tal vez nos permita considerar que trabajando con los maestros en formación inicial con el objetivo de que se posicionen en un perfil más próximo a un modelo indagador, sus ideas sobre la evaluación también cambiarán y viceversa. No obstante, para conocer la posible "bidireccionalidad" de esta relación es necesario realizar estudios en los que se trabaje en ambos sentidos: trabajando desde una evaluación cognitivista en sus formación y analizar la existencia de cambios en sus perfiles hacia modelos más indagadores o trabajar desde la indagación para evaluar los cambios en sus visiones sobre la evaluación.

Tal vez el modo en el que se aborda la formación inicial de maestros vaya en este último sentido, es decir, hacia la realización de actividades para promover un cambio hacia modelos más indagadores, y este sea el motivo por el que la evaluación no sea apenas considerada en estos estudiantes (Gil Quílez, Martínez Peña y Cordero, 2017). En palabras de Sanmartí (2007): "La evaluación es el motor de aprendizaje, ya que de ella depende tanto qué y cómo se enseña, como el qué y cómo se aprende" (p.19) y la formación inicial del profesorado no debería ser ajena a la misma.

Agradecimientos

Grupo Consolidado de Investigación Aplicada BEAGLE (Gobierno de Aragón y Fondo Social Europeo) perteneciente al Instituto Universitario de Ciencias Ambientales de Aragón (IUCA-UNIZAR). Proyecto EDU2016-76743-P del MINECO.

Referencias bibliográficas

- Abril, A.M., Ariza, M.R., Quesada, A., García, F.J. (2014). Creencias del profesorado en ejercicio y en formación sobre el aprendizaje por investigación. *Revista Eureka sobre enseñanza y divulgación de las Ciencias*, 11, 22-33. Recuperado de: <https://revistas.uca.es/index.php/eureka/article/view/2860>
- Alonso, A., Gil, D. y Martínez-Torregrosa, I. (1996). Concepciones docentes sobre la evaluación en la enseñanza de las ciencias. *Alambique*, 4, 6-15.
- Anderson, R. D. (2002). Reforming science teaching: what research says about inquiry? *Journal of Science Teacher Education*, 13(1), 1-12. DOI: <https://dx.doi.org/10.1023/A:1015171124982>

- Atkin, J.M, Black, P. y Coffey, J.E. (2001). *Classroom assessment and the National Science Education Standards*. Washington, DC: National Academy Press.
- Barnes, N., Fives, H. y Dacey, C.M. (2017). U.S. teachers' conceptions of the purposes of assessment. *Teaching and Teacher Education, 65*, 107-116. DOI: <https://dx.doi.org/10.1016/j.tate.2017.02.017>
- Brown, G. (2004). Teachers' conceptions of assessment: implications for policy and professional development. *Assessment in Education, 11*(3), 301-318.
- Brown, T.A. (2006). *Confirmatory factor analysis for applied research*. New York: The Guilford Press.
- Bryan, L.A. (2003). Nestedness of Beliefs: Examining a Prospective Elementary Teacher's Belief System about Science Teaching and Learning. *Journal of research in science teaching, 40* (9), 835-868.
- Bryan, L.A. (2012). Research on Science Teacher Beliefs. En B.J. Fraser, K.G. Tobin y C.J. McRobbie (Eds.), *Second International Handbook of Science Education* (pp. 477-495). Nueva York: Springer.
- Campanario, J.M. y Moya, A. (1999). ¿Cómo enseñar ciencias? Principales tendencias y propuestas. *Enseñanza de las Ciencias, 17*, 179-192.
- Campbell, T., Abd-Hamid, N.H. y Chapman, H. (2010). Development of instruments to assess teacher and student perceptions of inquiry experiences in science classrooms. *Journal of Science Teacher Education, 21*(1), 13-30. DOI: <https://dx.doi.org/10.1007/s10972-009-9151-x>
- Cañal, P., Travé, G. y Pozuelos, F. (2011). Análisis de obstáculos y dificultades de profesores y estudiantes en la utilización de enfoques de investigación escolar. *Investigación en la escuela, 73*, 5-26.
- Cortés, A.L., Gándara, M. de la, Calvo, J.M., Martínez, M.B., Ibarra, M., Arlegui, J. y Gil, M.J. (2012). Expectativas, necesidades y oportunidades de los maestros en formación ante la enseñanza de las Ciencias en la Educación Primaria. *Enseñanza de las Ciencias, 30* (3), pp. 155-176
- Couso, D. (2014) De la moda de "aprender indagando" a la indagación para modelizar: una reflexión crítica. En M.A. Heras, A. Lorca, B. Vázquez, A. Wamba, R. Jiménez. *Investigación y transferencia para una educación en ciencias: Un reto emocionante* (pp. 1-28). Huelva: Servicio de Publicaciones Universidad de Huelva.
- Flick, L.B. y Lederman, N.G. (2006). *Scientific Inquiry and Nature of Science: Implications for Teaching, Learning, and Teacher Education*. The Netherlands: Springer.
- Forbes, C.T. (2011). Preservice Elementary Teacher's Adaptation of Science Curriculum Materials for Inquiry- Based Elementary Science. *Science education, 95*, 927-955.
- García-Carmona, A. y Acevedo Díaz, J.A. (2016). Concepciones de estudiantes de profesorado de Educación Primaria sobre la naturaleza de la ciencia: Una evaluación diagnóstica a partir de reflexiones en equipo. *Revista mexicana de investigación educativa, 21* (69), 583-610.
- Garritz, A. (2014). Creencias de los profesores, su importancia y cómo obtenerlas. *Educación Química, 25*(2), 88-92.
- Gil Pérez, D. (1983). Tres paradigmas básicos en la enseñanza de las ciencias. *Enseñanza de las Ciencias, 1*, 26-33.

- Gil Quílez, M.J., Martínez Peña, M.B., De la Gándara Gómez, M., Calvo Hernández, J.M. y Cortés Gracia, A. (2008). De la universidad a la escuela: no es fácil la indagación científica. *Revista Interuniversitaria de Formación del Profesorado*, 63 (22,3), 81-100.
- Gil Quílez, M.J., Martínez, M.B. y Cordero, S. (2017). Grabaciones de situaciones de aula para la formación del profesorado. *Ápice. Revista de Educación Científica*, 1(1), 58-73. DOI: <https://doi.org/10.17979/arec.2017.1.1.2005>
- Hamilton, M. (2018). Pedagogical transitions among science teachers: how does context intersect with teacher beliefs? *Teachers and Teaching*, 24(2), 151-165. DOI: <https://doi.org/10.1080/13540602.2017.1367658>
- Hashweh, M.Z. (1996). Effects of science teachers' epistemological beliefs in teaching. *Journal of Research in Science Teaching*, 33(1), 47-63.
- Ibarra, M.S. y Rodríguez, G. (2010). Aproximación al discurso dominante sobre la evaluación del aprendizaje en la universidad. *Revista de Educación*, 351, 385-407.
- Jiménez Aleixandre, M.P. (2000). Modelos didácticos. En F.J. Perales y P. Cañal (eds.), *Didáctica de las ciencias experimentales. Teoría y práctica de la enseñanza de las ciencias* (pp. 165-186). Alcoy: Marfil.
- Kirschner, P.A., Sweller, J. y Clark, R.E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41, 75-86.
- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A. y Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada. *Anales de psicología*, 30, 1151-1169.
- Lucha López, P., Bravo Torija, B., Forcadell Aznar, L. y Ferrer Bueno, L. M. (2016). ¿Cuál es el modelo docente de los estudiantes de magisterio cuando desempeñan el rol de maestros? *Campo abierto: Revista de educación*, 35(1), 83-92.
- Markić, S., Eilks, I. y Valanides, N. (2008). Developing a Tool to Evaluate Differences in Beliefs About Science Teaching and Learning Among Freshman Science Student Teachers from Different Science Teaching Domains: A Case Study, *Eurasia Journal of Mathematics, Science & Technology Education*, 4(2), 109-120.
- Martín, C., Prieto, T. y Jiménez, M.A. (2013). Algunas creencias del profesorado de ciencias en formación sobre la enseñanza de la problemática de la energía. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 10 (Núm. Extraordinario), 649-663.
- McMillan, J.H., Myran, S. y Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *The Journal of Educational Research*, 95(4), 203-213.
- Muthén, L.K. y Muthén, B.O. (1998-2010). *Mplus User's Guide*. Sixth Edition. Los Angeles: Muthén y Muthén.
- NRC, National Research Council (2000). *Inquiry and the National Science Education Standards. A Guide for Teaching and Learning*. Washington, D.C.: National Academy Press.
- Pajares, F. (1992). Teacher's beliefs and education research: cleaning up a messy construct. *Review of Educational Research*, 62(3), 307-332.
- Pontes Pedrajas, A., Poyatos López, F. J. y Oliva Martínez, J. M. (2016). Concepciones sobre evaluación en la formación inicial del profesorado de Ciencias, Tecnología y Matemáticas. *Revista Iberoamericana de Evaluación Educativa*, 9(1), 91-107.

- Porlán, R. (1993). *Constructivismo y escuela: hacia un modelo de enseñanza-aprendizaje basado en la investigación*. Sevilla: Díada.
- Porlán, R. y Martín del Pozo, R. (2004). The conceptions of In-service and Prospective Primary School Teachers About the Teaching and Learning of Science. *Journal of Science Teacher Education*, 15(1), 39-62.
- Porlán, R. y Rivero, A. (1998). *El conocimiento de los profesores*. Sevilla: Díada Editora.
- Pozo, J.I. (2000). Concepciones de aprendizaje y cambio educativo. *Ensayos y Experiencias*, 33, 4-13
- Rivero, A., Martín del Pozo, R., Solís, E. y Porlán, R. (2017). *Didáctica de las Ciencias Experimentales en Educación Primaria*. Madrid: Editorial Síntesis.
- Rodríguez Barreiro, L.M., Gutiérrez Muzquiz, F.A. y Molledo, J. (1992). Una propuesta integral de evaluación en Ciencias. *Enseñanza de las Ciencias*, 10 (3), 254-267.
- Romero Ariza, M. (2017). El aprendizaje por indagación: ¿existen suficientes evidencias sobre sus beneficios en la enseñanza de las ciencias? *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 14(2), 286-299. Recuperado de: <https://revistas.uca.es/index.php/eureka/article/view/3335>
- Sáez Bondía, M.J. y Cortés Gracia, A.L., Gil Quílez, M.J., Martínez Peña, B., Orejudo Hernández, S. (2017). Perfiles de maestros en formación ante la enseñanza de las ciencias por indagación. *Enseñanza de las ciencias*, nº extraordinario, 2213-2218.
- Sampson, V. y Benton, A. (2006). Development and validation of the beliefs about reformed science teaching and learning (BARSTL) questionnaire. In *Annual Conference of the Association of Science Teacher Education (ASTE)*. Portland, Oregon.
- Sanmartí, N. (2007). *10 ideas clave. Evaluar para aprender*. Barcelona: Graó.
- Stiggins, R. (2006). What a difference a word makes. Assessment FOR learning rather than assessment OF learning helps students succeed. *Journal of Staff Development*, 27(1), 10-14.
- van Aalderen-Smeets, S.I., Walma van der Molen, J.H., van Hest, E.G.C. y Poortman, C. (2017). Primary teachers conducting inquiry projects: effects on attitudes towards teaching science and conducting inquiry. *International Journal of Science Education*, 39(2), 238-256. DOI: <https://doi.org/10.1080/09500693.2016.1277280>
- van Uum, M.S.J., Verhoeff, R.P. y Peeters, M. (2016). Inquiry-based science education: towards a pedagogical framework for primary school teachers. *International Journal of Science Education*, 38(3), 450-469. DOI: <https://doi.org/10.1080/09500693.2016.1147660>
- Wallace, C.S. y Kang, N. (2004). An investigation of Experienced Secondary Science Teachers' Beliefs About Inquiry: An Examination of competing belief sets. *Journal of research in science teaching*, 41(9), 936-960.
- Windschitl, M. (2003). Inquiry Projects in Science Teacher Education: What Can Investigative Experiences Reveal About Teacher Thinking and Eventual Classroom Practice? *Science Education*, 87, 112-143.
- Yoon, H.G. y Kim, B.S. (2016). Teachers' beliefs about nature of science and constructivist teaching in the content-specific context. *Eurasia Journal of Mathematics, Science & Technology Education*, 12(3), 457-475.

