



Una investigación sobre la evolución del conocimiento didáctico del profesorado sobre la evaluación en Ciencias

A piece of research on the evolution of teachers' didactic
knowledge about assessment in science

Lidia López-Lozano

Escuela Universitaria de Osuna, Universidad de Sevilla, Sevilla, España.

lidialopez@us.es

Emilio Solís Ramírez

Departamento de Didáctica de las Ciencias Experimentales y Sociales, Universidad de Sevilla, Sevilla, España.

esolis@us.es

RESUMEN: • Presentamos un trabajo para investigar los procesos de cambio en el conocimiento didáctico del profesorado en formación sobre la evaluación en Ciencias. Se aplica una metodología mixta bajo un diseño concurrente con técnicas cualitativas y cuantitativas utilizando un proceso de triangulación. Participan 347 estudiantes del Grado de Educación Primaria que cursan la asignatura de Didáctica de las Ciencias Experimentales. Analizamos cuestionarios, guiones formativos y los diseños elaborados por dichos estudiantes. El procedimiento seguido se muestra valioso y útil para documentar la evolución experimentada del profesorado. Los resultados evidencian cambios relativos a una mayor coherencia entre el conocimiento declarativo y factual del profesorado al finalizar la formación estimulado en gran medida por la reflexión sobre su diseño.

PALABRAS CLAVE: Formación inicial del profesorado; Métodos mixtos; Evaluación alternativa; Enseñanza de las ciencias.

ABSTRACT • We present a study to delve in the processes of change in pre-service teachers' didactic knowledge on assessment in science. A mixed methodology is applied under a concurrent design with qualitative and quantitative techniques using a triangulation process. The study involves 347 students from the degree in Primary Education taking the subject of Experimental Sciences Didactics. Questionnaires, instructional activities and designs of teaching units for data collection are applied. The procedure followed is valuable and useful for documenting the evolution in teachers' conceptions. The results show changes related to a greater coherence between the declarative and factual knowledge of the teachers at the end of the course, stimulated by the reflection on their practice.

KEYWORDS: Pre-service teacher education; Mixed methods; Alternative assessment; Science education.

Recepción: agosto 2018 • Aceptación: febrero 2019 • Publicación: marzo 2020

López-Lozano, L. y Solís Ramírez, E. (2020). Una investigación sobre la evolución del conocimiento didáctico del profesorado sobre la evaluación en Ciencias. *Enseñanza de las Ciencias*, 38(1), 87-104.

<https://doi.org/10.5565/rev/ensciencias.2755>

INTRODUCCIÓN

Este trabajo pretende mostrar el procedimiento seguido para investigar los procesos de cambio en el conocimiento didáctico del profesorado sobre la enseñanza de las ciencias y, específicamente, sobre la evaluación en ciencias. Esta aportación forma parte de los resultados obtenidos con el proyecto I+D+i,¹ cuyo propósito ha sido desarrollar y valorar un modelo de formación del profesorado basado en la investigación de problemas curriculares relevantes y en la interacción con prácticas docentes innovadoras desde la perspectiva constructivista del proceso de enseñanza-aprendizaje (Rivero, Martín del Pozo, Solís, Azcárate y Porlán, 2017).

En este trabajo presentamos los cambios que se ha detectado que se producen en las concepciones del profesorado sobre evaluación en ciencias como resultado de la formación recibida, haciendo una exposición detallada del desarrollo metodológico que hemos seguido, ya que pensamos que este nos permite aportar evidencias de estos cambios. Para ello, nuestra investigación se nutre de los avances y las aportaciones desarrollados en dos ámbitos disciplinares. Por una parte, la formación del profesorado en didáctica de las ciencias y, particularmente, su conocimiento didáctico sobre la evaluación en ciencias. Y, por otra, los desarrollos metodológicos en el ámbito de la investigación educativa, en especial las aportaciones más recientes sobre métodos mixtos de investigación en educación.

MARCO TEÓRICO

La evaluación en el proceso de enseñanza y aprendizaje en ciencias

Dada la complejidad de los procesos de enseñanza, se hace necesario entender el conocimiento que los profesores «manejan» en dichos procesos. Desde hace décadas, este ha sido objeto de numerosas investigaciones e interpretaciones. Por ejemplo, Friedrichsen, van Driel y Abell (2011), en relación con la conceptualización y el desarrollo del conocimiento profesional, plantean dos grandes enfoques: uno, centrado en el profesorado y, otro, los que se focalizan en el alumnado.

Este trabajo se nutre de los planteamientos del Proyecto IRES,² que distingue entre conocimiento profesional dominante y conocimiento profesional deseable, en función del grado de integración de los saberes implicados en el conocimiento profesional. En este sentido recurrimos a la idea de modelo didáctico como herramienta de análisis y actuación reflexiva durante la formación del profesorado, entendiéndolo como un elemento regulador y dialéctico entre lo que se piensa y lo que se hace (Solís y Porlán, 2003). El modelo didáctico atiende a los elementos curriculares y las concepciones epistemológicas más relevantes del profesorado (García, 2000).

Son varias las tipificaciones propuestas sobre los modelos didácticos, desde los contextos más genéricos hasta el ámbito de las ciencias, que es, concretamente, el que nos ocupa (véanse Porlán y Rivero, 1998; Sickel, 2017). No obstante, comparten el presentarlos como un continuo entre el modelo basado en la transmisión-recepción de conocimiento –el tradicional– y el constructivista –o alternativo o de investigación escolar, según IRES, nuestro marco de referencia–. De manera que, al analizar las concepciones de los profesores, el primero pueda actuar de «nivel de partida», fundamentalmente, influenciado por la tradición académica arraigada, en contraposición a aquel coherente con las investigaciones

1. Proyecto I+D+i EDU2011-23551: La progresión del conocimiento didáctico de los futuros maestros en un curso basado en la investigación y en la interacción con una enseñanza innovadora de las ciencias, financiado por el entonces Ministerio de Ciencia e Innovación.

2. Grupo Investigación en la Escuela (1991). *Proyecto curricular «Investigación y Renovación Escolar» (IRES)*. Introducción y cuatro volúmenes. Sevilla: Díada.

didácticas actuales basadas en procesos de investigación escolar (NGSS, 2013; Rocard et al., 2007), que sería el «nivel deseable», contemplando entre ambas situaciones intermedias (Porlán et al., 2010).

En el ámbito de la ciencia, desde la perspectiva de la alfabetización científica (Goodrum, Hackling y Rennie, 2001), las investigaciones en didáctica de las ciencias han puesto de relieve el fracaso de los modelos de enseñanza y aprendizaje tradicionales basados en la mera transmisión y recepción de conocimientos frente a la potencialidad de orientar la enseñanza de las ciencias mediante la investigación escolar (COSCE, 2011), lo que se denomina *Inquiry-based Science Education* (acrónimo, IBSE), que se puede considerar un referente en la enseñanza de las ciencias (NGSS, 2013).

Paralelamente, durante el siglo xx la evaluación experimenta una evolución en el ámbito educativo (Escudero, 2016; Shepard, 2000) acorde con los cambios hacia teorías de aprendizaje socioconstructivistas. Estas teorías enfatizan, sobre todo, la función que tienen durante todo el proceso, los criterios y la metodología que se debe seguir, de acuerdo con un currículo que prioriza la construcción de un conocimiento significativo en lugar de la memorización de contenido científico. En este escenario, la evaluación se presenta como un componente crítico del proceso de enseñanza y aprendizaje basado en la investigación (Harlen, 2013).

Tradicionalmente, se piensa en términos «de medición», referida esta al rendimiento de los escolares y centrada en la adquisición de los contenidos conceptuales. Esto viene siendo *grosso modo* la visión más «simplista». No obstante, en el proceso de evaluación, pensamos que lo que cobra importancia son las finalidades que lleva implícito: una responde a la ayuda o progresión del aprendizaje –formativa– y la otra informa de lo aprendido en un momento determinado –sumativa– (Bennet, 2011).

Así, se puede distinguir un modelo teórico bipolar en el continuo del proceso de evaluación según el objetivo que se persiga. Se distingue entre la «cultura del examen» (*testing culture*) y la «cultura de la valoración» (*assessment culture*) (Remesal, 2011). Se hace referencia a la evaluación como medida y la evaluación como investigación (Hargreaves, 2005). Se trata de una evaluación de aprendizaje frente a la evaluación para el aprendizaje (o evaluación formativa) (Black y Wiliam, 2009; Harlen, 2013). Estos dos polos se enmarcan en dos modelos o enfoques de enseñanza-aprendizaje. Una evaluación ligada a una enseñanza de la ciencia basada en la transmisión y recepción de los conocimientos, en contraposición a una evaluación enmarcada en una enseñanza constructivista basada en procesos de investigación escolar.

Desde esta óptica, concebimos la evaluación como «un proceso de investigación que intenta dar cuenta, permanentemente, del estado de evolución de las concepciones o ideas de los alumnos, de la actuación profesional del profesor y, en definitiva, del propio funcionamiento del proyecto de trabajo» (García, 2000).

El enfoque plurimetodológico en la investigación en formación del profesorado

En el terreno de la investigación educativa, la complejidad del sujeto del estudio es tal que, como indica Bisquerra (2009), se opta por abordar y conceptualizar su estudio desde diversas perspectivas. De acuerdo con esto, Albert (2007) propone una perspectiva empírico-analítica, una concepción interpretativa para comprender la conducta humana y así poder interpretar y comprender los fenómenos educativos más que aportar explicaciones de tipo causal. También se debe tener presente la corriente crítica, que trata de desvelar creencias, valores y supuestos que subyacen en la práctica educativa.

En nuestro estudio, la existencia de elementos y supuestos aceptables en todos los casos nos lleva a proponer una combinación adecuada de ellos, es decir, un enfoque plurimetodológico (Creswell y Plano, 2010; Tashakkori y Creswell, 2007). Sobre la integración de métodos, podemos decir que está ampliamente aceptada (Sánchez-Gómez, 2015). En este trabajo, esta combinación intenta potenciar la investigación armonizando lo cualitativo y lo cuantitativo, lo subjetivo y lo objetivo, lo empírico y lo interpretativo.

En el contexto de la formación inicial, para investigar acerca del conocimiento didáctico sobre el problema curricular de la evaluación, creemos que un enfoque mixto de investigación aporta unas conclusiones más potentes. En este sentido, la investigación complementa lo más característico de la tendencia interpretativa –conocer lo que ha ocurrido durante el proceso, la perspectiva de los participantes, el análisis documental–, con los aspectos relacionados con la metodología cuantitativa –contar con un número importante de participantes y usar la estadística en el análisis de ciertos datos–, siendo el núcleo central la perspectiva de los sujetos participantes en el proceso y trabajando desde la creencia en la posibilidad de *transformar* su conocimiento escolar y profesional mediante la acción, perspectiva sociocrítica. Respecto a este último enfoque, como se verá al describir el contexto de esta investigación, se comprenderá que «esta forma de hacer investigación tiene como constante ayudar a los participantes a utilizar la investigación como herramienta para desarrollar pensamientos, conductas críticas y comprender su entorno» (Colás, Buendía y Hernández, 2009, p. 121).

En la investigación utilizaremos, por tanto, una combinación de métodos cualitativos y cuantitativos teniendo como referente *el discurso de la integración y complementariedad metodológica* (Bisquerra, 2009, p. 78), de manera que la relación entre teoría y metodología o técnica sea lo más flexible posible. Siguiendo a este autor, la adopción de esta postura se ve reflejada, principalmente, en dos consecuencias prácticas. La primera es el uso conjunto de los diferentes métodos según las exigencias de la investigación, de manera que las decisiones se toman bajo una *perspectiva sociológica de la ciencia* (De Miguel, 1988, p. 62, citado en Bisquerra, 2009). La segunda atiende a la complejidad de los problemas educativos, que requiere contar con un pluralismo de enfoques que combinen datos, métodos y técnicas de investigación sin decantarse por una sola tendencia metodológica.

Compartimos con Ruiz-Olabuénaga (2012, p. 58) que en la actualidad «la mayoría de los autores, [...], adoptan posturas más funcionales pragmáticamente y menos reducibles teóricamente, siendo la corriente más numerosa la de aquellos que adoptan una postura irénica».

FINALIDAD

El propósito de este trabajo, pues, es estudiar la evolución del conocimiento didáctico del profesorado sobre la evaluación en ciencias a partir de un proceso de formación basado en la investigación y la reflexión sobre su práctica a la vez que proponemos un procedimiento metodológico válido y útil para ello. Esto, pensamos, nos va a permitir responder, por una parte, a la pregunta «¿Qué cambios se detectan en el conocimiento didáctico de los maestros de Primaria en formación sobre la evaluación en Ciencias cuando participan en un curso de orientación constructivista basada en la investigación escolar?». Y, por otra parte, a la cuestión «¿Qué grado de coherencia se observa entre sus creencias y concepciones sobre la evaluación y las propuestas didácticas que diseñan a partir de las reflexiones y análisis que hacen de estas?».

METODOLOGÍA

Contexto y participantes

Nuestra investigación tiene como contexto un curso formativo diseñado dentro del Proyecto I+D+i (–Cuaderno APENCIP–, Rivero et al., 2012) para la asignatura Didáctica de las Ciencias Experimentales (9 créditos) correspondiente al 2.º curso del Grado de Educación Primaria de la Facultad de Ciencias de la Educación de la Universidad de Sevilla. Se investigaron cinco clases dirigidas por cinco formadores distintos que compartían el mismo programa formativo, durante el curso 2012-2013, implicando a un total de 347 estudiantes agrupados en 92 equipos de trabajo de 3 a 6 estudiantes. Las

clases presentan características demográficas y académicas similares. La mayoría (63,4 %) son mujeres, con edades en torno a los 20 años, que cursan por primera vez esta asignatura y presentan escasa formación en ciencias.

El curso se fundamenta en la investigación profesional de problemas curriculares relevantes –contenidos escolares, ideas de los alumnos, metodología y evaluación– y en la interacción con prácticas docentes innovadoras (mediante audiovisuales y casos prácticos). La secuencia de actividades se orienta bajo un enfoque coincidente con lo que Abell, Appelton y Hanuscin (2010) denominan *Reflection Orientation* (Orientación de la reflexión). De modo que este aprendizaje requiere la evaluación y reformulación de sus propias teorías, contrastándolas con nuevas informaciones (documentos de apoyo, declaraciones y prácticas reales de maestros cuando hacen investigación escolar en sus aulas). Se le plantea un supuesto práctico: *el diseño de una propuesta para enseñar a alumnos de Primaria un contenido del área de Ciencias de la Naturaleza*. En Rivero, Martín del Pozo, Solís y Porlán (2017) se describe ampliamente el proceso con todas sus características y sus apoyos teóricos. No obstante, la secuencia formativa se puede resumir como sigue:

Se selecciona, por parte de cada equipo, un contenido del currículo de Ciencias de Primaria para elaborar la primera propuesta didáctica partiendo de sus conocimientos y experiencias (Diseño1, DS1). A continuación, iniciamos un ciclo de trabajo para cada uno de los problemas curriculares que se van a investigar y que contempla: el análisis de lo elaborado –guion de análisis (GA)–, el contraste –documentos (escritos y audiovisuales) y puestas en común– y la reflexión –guion de reflexión (GR)–. Tras esto, se pasa a la elaboración de la segunda propuesta de enseñanza (Diseño2, DS2), que recoge las modificaciones que consideran oportunas. Como última actividad se intenta acercar a los estudiantes a una práctica real utilizando vídeos con secuencias completas de actividades en aulas donde se lleva a cabo una enseñanza de las ciencias basada en la investigación escolar (Ezquerro y Rodríguez-Marín, 2013), debatir sobre ello y cumplimentar un guion sobre estas prácticas observadas (GP). Esto conduce a la elaboración de una tercera y definitiva propuesta (Diseño 3, DS3), que se compara con las dos primeras. Por lo tanto, la construcción progresiva del conocimiento está muy presente mediante la reelaboración continua de la propuesta didáctica que diseñan los equipos (figura 1).

Además, tanto al inicio como al final del curso los estudiantes cumplimentan un cuestionario tipo Likert sobre el aprendizaje y la enseñanza de las ciencias, con el objetivo de tener una referencia sobre su orientación o modelo didáctico en ambos momentos del curso, lo que permite contrastar si se ha producido algún tipo de cambio.

En este trabajo nos centramos en la parte de la secuencia formativa (pretest-diseño-análisis-contraste-reflexión-diseño-postest) correspondiente a la investigación sobre la evolución del conocimiento didáctico de la evaluación en ciencias (figura 1).

Instrumentos de recogida de datos

De acuerdo con lo anteriormente expuesto, el propio desarrollo del proceso didáctico, escenario de nuestra investigación, nos ha permitido contar con una gran variedad de instrumentos de recogida de información producidos en diferentes momentos de este (al inicio, M1; momento intermedio, M2; y al final, M3).

Por un lado, disponemos de los documentos elaborados por los estudiantes durante el curso, que nos permiten distinguir principalmente dos dimensiones. La primera dimensión (DM1) cuenta con los tres diseños (DS1, DS2 y DS3). Son documentos de estilo libre de escritura. Esta dimensión tiene un carácter dinámico porque nos informa de los posibles cambios de los futuros docentes en la realización de los diseños progresivos a lo largo del curso. Su análisis es de diseño proposicional. La segunda dimensión (DM2) recoge los guiones de análisis, de reflexión y sobre la práctica profesional

(GA, GR y GP). Permite un análisis a nivel reflexivo. Se analizaron un total de 270 diseños y 225 guiones. Por otro lado, disponemos del cuestionario de escala tipo Likert de 6 valores de cumplimentación individual (311 participantes), al principio y al final del curso (pretest y postest, CLikert_i y CLickert_p respectivamente) (figura 1). Elaborado *ad hoc* para el proyecto, fundamentado en estudios anteriores y validado por expertos del área (Hamed, 2013), en él los estudiantes expresan su nivel de acuerdo o desacuerdo respecto a diferentes afirmaciones sobre problemas curriculares, entre ellos la evaluación (objeto de estudio), y así se configura en 48 ítems en total, 12 para cada uno. En relación con la evaluación, se declara *para qué, qué y cómo evaluar* (en correspondencia con lo abordado en clase), y sus enunciados responden a un enfoque de evaluación tradicional (6 ítems) y a un enfoque de referencia (6 ítems) (López-Lozano, Solís y Azcárate, 2018). Esto nos da información desde el punto de vista declarativo.

Esta diversidad de instrumentos nos permite dar validez o, más apropiadamente en términos cualitativos, ofrecer credibilidad mediante la triangulación (Sánchez-Gómez, 2015).

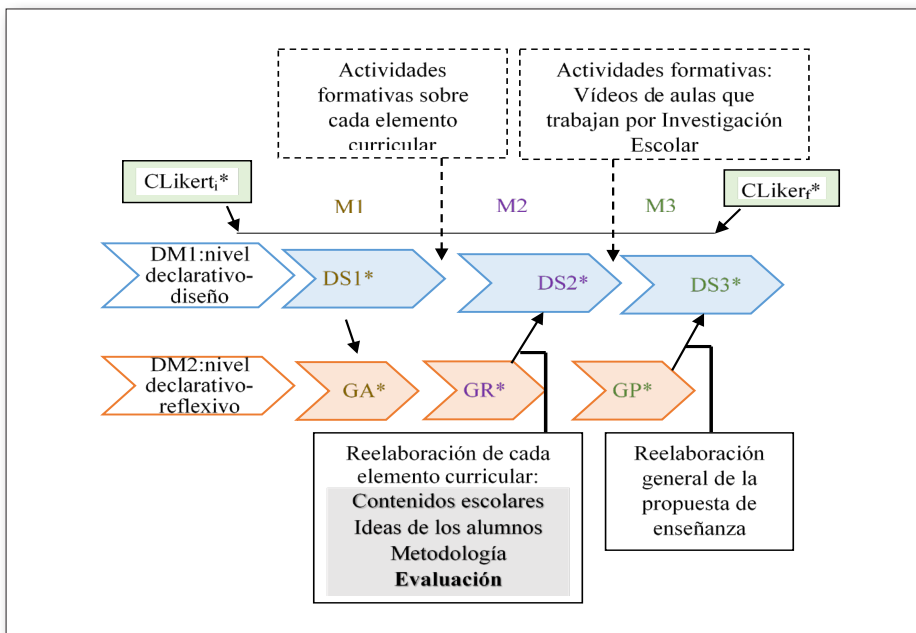


Fig. 1. Esquema del proceso formativo. Relación entre dimensiones (DM)-momentos (M)-instrumentos de recogida de datos (*). Adaptado de López-Lozano (2017).

Análisis de datos

Para el análisis de los documentos nos hemos basado en las técnicas de análisis de contenido (Bardin, 1996). Este análisis se entiende, tal y como expresa Ruiz-Olabuénaga (2012, p. 197), como «escenario de observación», de tal manera que de este clasificamos y codificamos los diferentes significados manifestados en categorías que representen, como indica Tójar (2006), el sentido de dicha manifestación escrita.

En nuestro caso, el proceso para definir nuestro sistema categorial como herramienta de análisis ha sido bidireccional (Osses, Sánchez y Ibáñez, 2006; Ruiz-Olabuénaga, 2012). Esto es, de la teoría a los datos y viceversa.

Dicho sistema se estructura en dos componentes. Por un lado, las categorías que sintetizan, conceptualmente, los aspectos que caracterizan la evaluación. La otra componente son los valores o

significados que adquieren dichas categorías, dependiendo de la posición teórica adoptada que, en nuestro caso, responde a la idea de progresión de aprendizaje (*Learning Progression*) (Dustchl, Maeng y Sezen, 2011), o bien itinerarios de progresión (IP en adelante) (Porlán et al., 2010). Así, hablamos de niveles de formulación para cada categoría, desde el de partida al deseable, vinculados a cada posición teórica, tradicional y alternativo, respectivamente. En lo que sigue se detalla el desarrollo en la configuración del sistema categorial debido al análisis bidireccional hasta llegar al sistema definitivo.

Se inició configurando un sistema de tres categorías y se formularon para cada una tres niveles de complejidad creciente, basándonos en resultados alcanzados en estudios anteriores del equipo de investigación del proyecto en otros ámbitos curriculares (Martín del Pozo, Porlán y Rivero, 2011; Martín del Pozo, Rivero y Azcárate, 2014; Rivero et al., 2011). Adaptando este sistema a la evaluación, consideramos que «el para qué» (categoría 1: sentido de la evaluación), «el cómo» (categoría 2: instrumentos de evaluación) y «el qué evaluar» (categoría 3: contenido que hay que evaluar) representan los aspectos centrales de la evaluación (López-Lozano, Solís y Fernández-Arroyo, 2018). Igualmente, estas constituyen las tres categorías contempladas en el cuestionario (López-Lozano, Solís y Azcárate, 2018). Cada una de ellas toma diferente significado según el enfoque evaluativo: desde el más tradicional (evaluación de aprendizajes) hasta el centrado en la evaluación para el aprendizaje (formativa y formadora) (Black y Wiliam, 2009).

Así, en lo relativo a los tres niveles previstos, el primer planteamiento constituye el nivel de partida mayoritario en el proceso de formación (Wang, Kao y Lin, 2010) y lo hemos denominado nivel 1 (N1, en adelante). Su formulación está vinculada a un enfoque tradicional de la evaluación y a una enseñanza caracterizada por la transmisión de conocimiento. En el otro extremo, el formativo, se halla el enfoque que consideramos de referencia, dado el consenso actual que existe entre colectivos de profesionales, investigadores y en las reformas educativas en la Enseñanza de las Ciencias, basado en la investigación escolar (Harlen, 2013; NGSS, 2013) y denominado nivel 3 (N3, en adelante). Representa un conocimiento de mayor complejidad, vinculado a un modelo didáctico de investigación en la escuela (MIE) (García, 2000).

Entre ambos niveles de conocimiento, de acuerdo con la idea de transición adoptada en los IP, existen enfoques intermedios que representan la situación más aceptada por los futuros maestros tras un proceso formativo. Esto es apoyado por los resultados obtenidos en los estudios citados anteriormente, en los que viene trabajando el equipo de investigación del proyecto sobre la evolución de conocimiento didáctico en docentes en formación. Así, establecemos, a modo de hipótesis en el caso de la evaluación, la existencia de un nivel de transición denominado nivel 2 (N2 en adelante). Este nivel adopta cierto carácter híbrido porque, por un lado, se formula a partir de la idea de transición en los procesos de cambio que contempla niveles de formulación intermedios y, por otro, la propuesta de un nivel de transición (N2) para evaluación coherente con los N2 detectados en el resto de los ámbitos curriculares de anteriores estudios y del propio proyecto.

En la figura 2 se sintetiza la configuración del sistema categorial inicial.

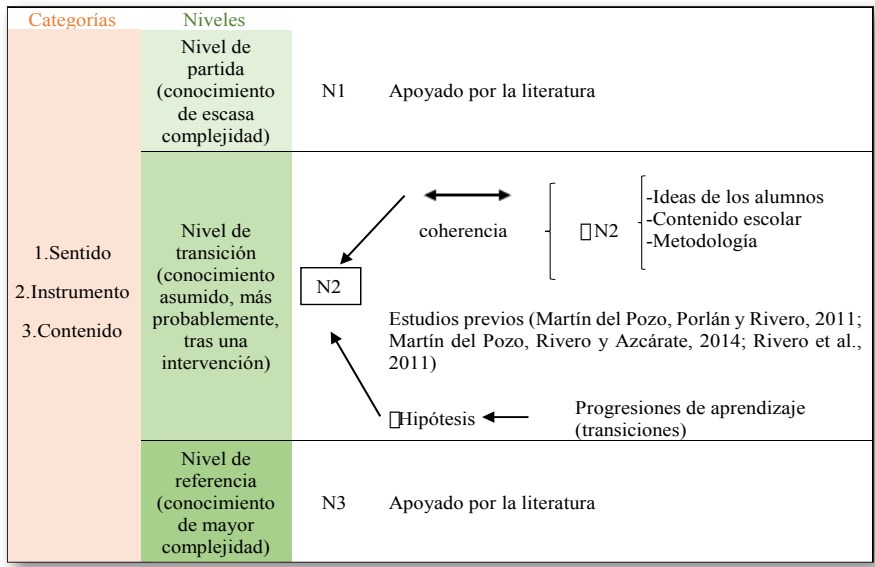


Fig. 2. Sistema de categorías inicial para el análisis de los documentos. Fuente: López-Lozano (2017).

Este sistema de tres categorías y niveles fue reformulado a lo largo del desarrollo del trabajo, ya que el proceso metodológico seguido «no es lineal, sino que es emergente y en cascada» (Colás et al., 2009, p. 100) hasta configurarse tal y como se muestra en la figura 3, mediante una constante negociación entre los datos y el proceso de análisis.

Efectivamente, emergieron de los datos informaciones (unidades de información de tipo tema, Bardin, 1996) que no era posible categorizar bajo los niveles establecidos; correspondían a situaciones intermedias entre los niveles N1 y N2 y los niveles N2 y N3, respectivamente. Son N1-2 (nivel uno, dos) y N2-3 (nivel dos, tres), por lo que fueron considerados, compartiendo con Smith, Wiser y Carraher (2010) (citado en Dustch et al., 2011) la importancia que tienen los niveles intermedios dentro de los procesos de aprendizaje y siendo conscientes de que el cambio requiere la superación de obstáculos que hacen que este sea gradual y paulatino. De esta forma, cada uno de los «escalones» o niveles intermedios forman un continuo que nos informa de dicho cambio (Porlán et al., 2010; Rodríguez-Marín, Fernández-Arroyo y García-Díaz, 2014). De igual modo, se detectó, sobre todo en el análisis de los primeros diseños, que no se incluía una propuesta de evaluación, con lo que definimos un nivel cero (N0) para indicar estos casos. Paralelamente, y como se ve en la figura 3, se reformulaban los enunciados de los niveles de cada una de las categorías según los datos y las consultas de la literatura pertinente.

ENFOQUE	<i>Categorías de estudio</i>			ANÁLISIS	
Tradicional (más o menos evolucionado)	N0	Sin dato			Inductivo
	Nivel	1-Sentido	2-Instrumento	3-Contenido	
	N0	Sin dato			Inductivo
	N1	Comprobar el nivel conceptual alcanzado por el alumnado respecto a los objetivos previstos. Medir niveles de aprendizaje.	El instrumento fundamental es el examen o prueba escrita.	El nivel conceptual alcanzado por el alumnado respecto a los objetivos previstos. Resultado.	Deductivo
	N1-2	Comprobar el nivel alcanzado por el alumnado, pero de alguna manera, también, sirve para valorar la enseñanza en función de los resultados obtenidos.	Además del examen, se consideran algunos otros instrumentos que se utilizan, puntualmente, a lo largo de la enseñanza.	El nivel conceptual principalmente, considerando de manera puntual el de procedimientos y/o el de actitudes alcanzado por el alumnado según los objetivos previstos.	Inductivo
De transición hacia el formativo (más o menos evolucionado)	N2	Valorar el cambio entre las ideas iniciales y finales del alumnado. Además, de la instrucción.	Diversidad de instrumentos para evaluar al alumno.	Nivel conceptual, procedimental y actitudinal. Además, algún aspecto de la propuesta de enseñanza.	Nivel híbrido deductivo-inductivo
	N2-3		Diversidad de instrumentos para evaluar al alumnado y, puntualmente, la instrucción.	Desarrollo de las capacidades y/o adquisición de competencias. Además, la actividad docente o la propuesta de enseñanza.	Inductivo
Modelo de referencia: formativo	N3	Regular y mejorar el proceso de enseñanza y aprendizaje.	Diversidad de instrumentos para evaluar el proceso de enseñanza y aprendizaje.	Progresión en el aprendizaje y la adecuación de la enseñanza para provocar dicho aprendizaje.	Deductivo
<i>Sistema de categorías</i>					

Fig. 3. Representación del diseño del sistema de categorías para el análisis de los datos. Fuente: Adaptado de López-Lozano (2017).

Estos niveles de formulación se pueden vincular a modelos de enseñanza; los niveles más simples se relacionan con un enfoque tradicional de la enseñanza y con una evaluación tradicional más o menos evolucionada (N0, N1 y N1-2). En estos niveles, el sentido de la evaluación es necesario, fundamentalmente para decidir sobre la promoción del alumno. Esta será positiva cuando el alumnado sea capaz de responder correctamente a las cuestiones que le plantea el profesorado y el nivel que deben alcanzar los alumnos. En el momento de la evaluación es el fijado en el programa de la asignatura. Si nos fijamos en el contenido de la evaluación, debe centrarse en los aprendizajes científicos esenciales, los relacionados con la comprensión de los conceptos, y en medir el nivel alcanzado respecto a los objetivos previstos. Si se hace referencia a los instrumentos de evaluación, el básico y más fiable para la evaluación de los

aprendizajes es el examen escrito. El nivel de mayor complejidad de conocimiento se relaciona con un enfoque de investigación escolar cuya evaluación es formativa y formadora (N3). En este nivel, el sentido de la evaluación debe preocuparse tanto del aprendizaje como de la enseñanza y ha de ser un instrumento básico para comprender y mejorar los procesos de enseñanza-aprendizaje. En relación con los contenidos, se debe considerar el aprendizaje de procedimientos y actitudes, además del de conceptos, así como la adecuación de la enseñanza para facilitar dicho aprendizaje. Finalmente, y con respecto a los instrumentos de evaluación, ha de utilizarse el máximo número de instrumentos posibles (cuadernos de clase, registros de participación, trabajo en el laboratorio, informes de autoevaluación, etc.) y preparar instrumentos de evaluación para evaluar a los alumnos, al profesor y la enseñanza desarrollada. Entre ambos están los niveles de transición hacia un enfoque formativo más o menos evolucionado (N2 y N2-3), en los que la evaluación mantiene la función de «medida», pero empieza a centrarse en el alumno e, incluso, en el proceso, e incorpora algunos instrumentos de evaluación, como podrían ser la observación, el análisis de los trabajos realizados y los ejercicios. Así, realmente sigue siendo una evaluación finalista, pero con cierto carácter procesual, aunque asistemática (López-Lozano, Solís y Fernández-Arroyo, 2018). Se trata de una evaluación a la que hemos llamado «transaccional» en el sentido de que se presenta como un enfoque que no deja de compartir, en cierta medida, una postura tradicional de la evaluación, pero con aproximaciones al enfoque formativo, dándose una especie de «transacción» conceptual entre ambos polos.

Concluido el proceso de categorización y codificación de las citas en cada categoría y nivel, se pasa a una fase de lectura global que nos permita asignar un nivel «definitivo» de formulación (un nivel global) a cada categoría en el momento formativo y en el documento correspondiente, lo que permite efectuar un recuento de las frecuencias de cada código (porcentajes) conociendo el peso de las categorías.

Por otra parte, la información que procede del cuestionario se somete a un tratamiento cuantitativo. Se trata de un diseño preexperimental de pretest/postest con un solo grupo. Se realizó un análisis descriptivo de los datos según los dos momentos de cumplimentación. Posteriormente, para determinar si existen diferencias significativas entre el pre y el postest, los datos fueron contrastados a través de la prueba inferencial no paramétrica de Mann-Whitney (U). Completando el contraste, realizamos el cálculo del tamaño del efecto para cuantificar la magnitud de las diferencias encontradas (Morales, 2007). Además, se constató la consistencia interna (Morales, 2012) y se comprobó empíricamente la estructura teórica de la escala propuesta (la validación de constructo) (Beavers et al., 2013), resultando que el constructo *concepciones sobre la evaluación en la enseñanza y el aprendizaje de las ciencias* presenta dos factores que comprenden cada uno el correspondiente conjunto de ítems elaborados considerando dos modelos teóricos, el tradicional y el formativo, respectivamente. El estudio estadístico se desarrolla detalladamente en López-Lozano, Solís y Azcárate (2018).

Una vez analizados los textos y las respuestas dadas al cuestionario (pre/postest), procedemos a realizar la triangulación de los datos, cuyos resultados se muestran a continuación.

RESULTADOS

En relación con los cambios producidos y el grado de coherencia de estos, podemos observar que hay cierta asimetría, fundamentalmente al inicio, entre lo que declaran al cumplimentar el cuestionario (el nivel identificativo) y lo que proponen en los diseños (el nivel declarativo-diseño), y entre esto y lo declarado en los guiones cuando analizan o reflexionan sobre sus planteamientos (el declarativo-reflexivo).

Si consultamos los datos estadísticos globales (figura 4), visto desde el conjunto de ítems de los dos enfoques de evaluación recogidos en el cuestionario —el tradicional (FET) y el formativo (FEF)—, vemos que, al inicio, las posturas hacia estos dos enfoques están bastante próximas a juzgar por la escasa

diferencia (0,97) en el promedio de las puntuaciones otorgada a ambos grupos de declaraciones. Así, estos futuros maestros se identifican fundamentalmente con ideas formativas, pero sin desprenderse de ciertas ideas tradicionales, mientras que al diseñar presentan, mayoritariamente, propuestas de evaluación de corte tradicional (el 93,3 % de las propuestas, niveles de formulación N0, N1 y N1-2). Iniciaban su propuesta manifestando una concepción de la evaluación como una mera constatación de si los alumnos son capaces o no de reproducir los conocimientos (40,21 %, N1). Es digno de resaltar que la mitad de las propuestas no manifestaban el sentido o la finalidad de la evaluación que se llevaría a cabo (50 %, N0). Básicamente, se evalúa a través del examen, ya sea en exclusividad (35,87 % N1) o acompañado de algún que otro instrumento (30,43 %, N1-2). No aportan información sobre lo que se pretende evaluar (34,78 %, N0), y cuando lo hacen, prevalece el conocimiento conceptual adquirido por el alumno (el 54,35 % de propuestas corresponden al N1 y N1-2). Léanse algunos ejemplos correspondientes al DS1:

Para finalizar el temario realizaremos un control de evaluación que nos ayudará a conocer lo que ha aprendido cada alumno sobre los climas de España [Equipo E18-EV1.N1; EV2.N1; EV3.N1].

EVALUACIÓN FINAL: Al finalizar cada tema evaluaremos a los niños y niñas de diversas maneras, es decir, no siempre se hará por escrito, también lo haremos de manera oral, y o a través de un pequeño trabajo en el aula (por ejemplo, juego de rol) [EV2.N1-2]. Pretendemos evaluar si han adquirido los objetivos esperados [EV1.N1]-Equipo J10.

En este momento inicial, también, cuatro equipos elaboran una propuesta acorde con una evaluación transaccional y dos equipos presentan una evaluación bajo el enfoque formativo de referencia, léase:

La evaluación es una acción que no debe centrarse únicamente en los alumnos (visión tradicionalista), sino alcanzar a todos los procesos y agentes implicados en la acción educativa [EV3.N3]. De este modo nuestra evaluación tendrá un carácter globalizador, formativo (mejora de los procesos), contextualizado, adaptándose a una realidad en concreto: nuestra aula y nuestros alumnos, así como de detección de posibles necesidades. Evaluamos para la toma de decisiones; para ir mejorando constantemente nuestra actuación (recordemos el carácter flexible que hemos dado a nuestro diseño-unidad didáctica) [EV1.N3]-Equipo F13.

A la hora de analizar sus primeros diseños, aproximadamente el 70 % se autoanalizan de acuerdo con un nivel de conocimiento más completo o complejo de lo que realmente han sido capaces de mostrar en el diseño.

Consecuentemente, en los diseños iniciales no se ve reflejado el alto nivel de acuerdo inicial expuesto con las declaraciones formativas (todo lo contrario), pero sí son más coherentes con lo manifestado acerca de las tradicionales. Así, podemos apreciar incoherencias si relacionamos, a grandes rasgos, lo diseñado con lo declarado en el cuestionario, y no tanto entre lo expresado en el cuestionario y el guion de análisis. Es decir, asimetría entre lo factual (lo propuesto) y lo declarativo y simetría en esto y lo reflexivo.

Entre la fase inicial e intermedia del curso es donde se producen los cambios más significativos en las propuestas; no así entre la fase intermedia y final, donde la tendencia mayoritaria (65,5 % de las propuestas) es mantener la misma propuesta del DS2 en el diseño final DS3 (DS2 \cong DS3).

Tras la estrategia formativa, la muestra está dividida entre diseños de corte tradicional (50,5 %) y de enfoque transaccional (48,4 %), lo que evidencia cambios en la tendencia de diseño, ya que empiezan a considerar la evaluación como vía para conocer cómo funciona la metodología, normalmente, según los resultados obtenidos de los alumnos (muchos aprobados implica una metodología correcta, y si no es así algo no ha sido adecuado) y una evaluación cuya finalidad es la de poder valorar el cambio del conocimiento de los alumnos al final del proceso respecto a las ideas iniciales evaluadas –que sabían los

alumnos sobre el tema y qué acaban sabiendo– (35,9 %, N2), siendo esta concepción la más asumida, como se puede ver a continuación:

Una vez terminada la sesión realizaremos una actividad individual para conocer si los alumnos han adquirido nuevos conocimientos o si han modificado las ideas previas iniciales [Equipo A14-EV1.N2].

El uso predominante del examen pierde peso y las propuestas de instrumentos para evaluar al alumno se diversifican (50 %, N2), pero son anecdóticas aquellas que se preocupan de valorar el proceso de enseñanza en su conjunto:

*Actividades de evaluación:

- Identificar en un dibujo las partes del aparato circulatorio.
- Hacer mapas conceptuales de las funciones del corazón y la sangre.
- Que elaboren experimentos sobre los latidos del corazón.
- Trabajos donde busquen información de la actualidad como las enfermedades.
- Y por último un examen escrito del tema expuesto [Equipo C5-EV2.N2].

Pasamos de una evaluación centrada en los conocimientos teóricos a una en la que va tomando cierta presencia la valoración de las actitudes y las destrezas (36,95 %, N1-2 y 35,87 % N2):

No solo tendremos en cuenta una prueba final, sino también la actitud, el esfuerzo, la participación que el alumnado muestra durante todo el proceso [Equipo A14-EV3.N1-2].

Respecto al inicio, ha bajado casi a la mitad el número de diseños tradicionales, y esto coincide con las ideas que se reflejan en los guiones, ya que al reflexionar aparecen ideas más evolucionadas que en las propuestas (en torno al 85 y 74 % de los GR y GP, respectivamente, se muestran coincidentes o más complejos que en los últimos diseños). También, según los resultados del cuestionario (figura 4), las posturas tradicionales y formativas parecen alejarse viendo que la diferencia en el promedio es algo más del doble respecto a la situación inicial (2,02 puntos), de lo que resulta un gran acuerdo en las declaraciones formativas ($M = 5,14$) e indecisión tendente al desacuerdo en las tradicionales ($M = 3,12$). El cambio más decisivo, desde el punto de vista de la identificación, se ha dado en aquellos aspectos referentes al enfoque tradicional (a juzgar por el valor del tamaño del efecto, $d_{FET} = 0,58$), disminuyendo la incertidumbre mostrada con estos posicionamientos.

La mayoría se identifica con una evaluación que debe considerar no solo el aprendizaje conceptual, sino también el procedimental y el actitudinal (85,3 % en acuerdo, además, nadie en total desacuerdo), utilizándose el máximo número posible de instrumentos (82,9 % en acuerdo; además, nadie en total desacuerdo) con la finalidad de preocuparnos tanto del aprendizaje como de la enseñanza (81,6 % de acuerdo). Sin embargo, esto no ha significado, necesariamente, posicionarse en desacuerdo del todo con los posicionamientos tradicionales. Respecto a esto, encontramos que alejarse de una evaluación tradicional se relaciona también en los documentos con el rechazo del uso del examen.

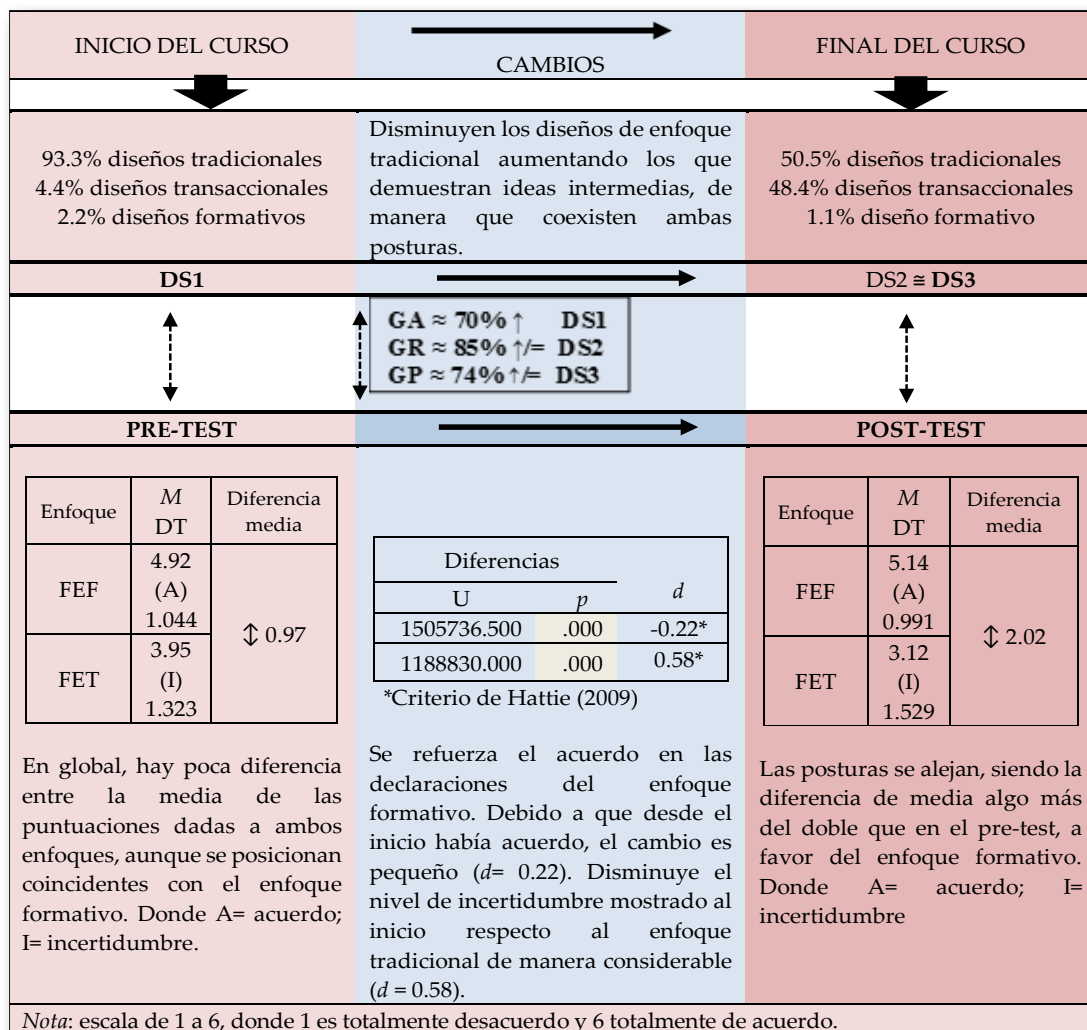


Fig. 4. Relación de los resultados globales obtenidos de los diseños, cuestionarios y guiones. Nota: FEF: factor de evaluación formativa; FET: factor de evaluación tradicional; DS1: diseño inicial; DS2: diseño intermedio; DS3: diseño final; GA: guion análisis; GR: guion reflexión; GP: guion de práctica. Fuente: Adaptado de López-Lozano (2017).

DISCUSIÓN Y CONCLUSIONES

Tal y como hemos ido desarrollando a lo largo de este trabajo, la discusión y las conclusiones que podemos plantear se pueden hacer, al menos, desde dos perspectivas. Por una parte, lo relacionado con los resultados de la investigación en lo que se refiere a los cambios que se han producido en el conocimiento sobre la evaluación de los sujetos de la muestra y, por otra, lo relacionado con el desarrollo metodológico de la investigación.

Respecto al conocimiento didáctico sobre la evaluación, concluimos que, al inicio del curso, hay cierta incongruencia entre lo que se propone y lo que responde a su deseo, plasmada en sus declaraciones y guiones, mostrándose de acuerdo con una evaluación favorecedora de aprendizajes, pero diseñando según una evaluación más aproximada a sus referencias como estudiantes (Buck et al., 2010; Wang et al., 2010). Tras la estrategia formativa, estas dimensiones (lo factual y lo declarativo) van acercándose, en lo que creemos que ha tenido un papel fundamental el hecho de explicitar, contrastar

y reflexionar sobre sus propias ideas, trasladando alguna de estas a su planificación final (Abell et al., 2010; Buck et al., 2010). En la evaluación en ciencias se aprecia una brecha entre lo que parece ser lo deseable (un componente más del proceso de enseñanza y aprendizaje que ayude al desarrollo intelectual del alumnado) y lo que «actúa» para no comprometer el rol del maestro (es decir, el elemento que nos permite mantener el control del aula) y en el papel de constatación del nivel alcanzado frente a la comunidad (Sanmartí, 2012). Esta «competitividad» entre lo que *se cree* y *se hace* se refleja, en cierta medida, en la asimetría presentada en el estudio. Fruto de este «acercamiento» final es la importante presencia del enfoque transaccional cuando se diseña.

Asimismo, concluimos que, a lo largo del curso, se va complejizando el conocimiento acerca de la evaluación en ciencias, pero el cambio se produce de manera gradual y algo compartimentada (Maclellan, 2004; Porlán et al., 2010; Solís y Porlán, 2003). Esto lo revela tanto el cuestionario como el análisis de los diseños, apoyado con el cruce de las manifestaciones volcadas en los guiones. Aun manifestando ideas formativas sobre la finalidad que debe adquirir la evaluación cuando reflexionan sobre ello tras el trabajo en clase, no han sido lo suficientemente potentes como para poder diseñar propuestas bajo esta premisa (1,1 % diseños formativos). En cambio, por lo general, las manifestaciones planteadas sobre la planificación de los instrumentos y del contenido eran bastante fieles en sus diseños. Acerca de la finalidad, se comprueba que ambos enfoques no se entienden como competitivos, por lo que resulta mucho más difícil el posicionamiento hacia una evaluación verdaderamente reflexiva y relacional destinada a la mejora y a la regulación de los procesos. Sobre esta categoría, en el estudio de Rivero et al. (2017), en el que se aborda el cambio del conocimiento sobre la enseñanza de las ciencias en la misma muestra de sujetos (92 equipos, las propuestas de diseño), analizándolo junto a la presentación de los contenidos a los alumnos, la utilización didáctica de sus ideas y la secuencia metodológica seguida, se obtiene que esta finalidad de la evaluación es la más resistente al cambio. Estos resultados, y de acuerdo también con Rivero et al. (2017), nos hacen pensar en la necesidad de tener muy presente, a la hora de formular y plantear una propuesta formativa, la obligación de incidir en aquellos aspectos, como es este del sentido de la evaluación, que parece que generan más dificultades para el cambio. Esto implicaría, por ejemplo, tener procesos formativos de ida y vuelta, con suficiente capacidad de provocar la reflexión en los maestros en formación y favorecer la retroalimentación de una manera continua con estos.

Por otra parte, y en lo que se refiere a las conclusiones relacionadas con la metodología de la investigación, que hemos desarrollado ampliamente, consideramos que una de las virtualidades de esta investigación ha sido el tamaño de la muestra. Hemos contado con 92 equipos de trabajo, 495 documentos elaborados por estudiantes-maestro y tener referencias de más de 300 estudiantes.

Por otra parte, haber dispuesto de distintos instrumentos de recogida de datos y en distintos momentos de la formación, combinando así una fase cuantitativa con otra más cualitativa, ha dado como resultado tres puntos de datos, de distinto nivel de análisis, combinando lo identificativo con lo declarativo y lo reflexivo. Creemos que esto nos permite alcanzar conclusiones más fundamentadas.

Al hilo de lo apuntado, debemos tener presente que este cruce de datos, provenientes de instrumentos con distinto tratamiento metodológico implica varios cambios de «lente» con los que mirarlos (Ruiz-Olabuénaga, 2012): cambio de objeto, cambio de lenguaje (de lo cualitativo a lo cuantitativo y viceversa, así pasamos de números a constructos teóricos y de conceptos a números) y cambio de nivel (de lo declarativo a lo reflexivo y a lo proposicional).

Conforme al tratamiento de los documentos, el sistema de categorías inicial basado en hipótesis teóricas y estudios previos se ha visto enriquecido con niveles de formulación no previstos (los niveles N0, N1-2 y N2-3). Esto se debe a que ha sido un proceso dinámico y de constante negociación entre los fundamentos teóricos en los que se apoya y las variaciones que las informaciones de los datos fueron aportando. Como consecuencia, podemos proponer un sistema más fundamentado en datos empíricos.

En cuanto al cuestionario, los resultados del análisis factorial nos permiten hablar en términos de enfoques de evaluación –el tradicional y el formativo–, lo que nos lleva a desarrollar el estudio descriptivo e inferencial tomando dos grupos de ítems según perspectivas. Esto ha propiciado unificar el lenguaje en la interpretación de los diferentes datos.

Finalmente, creemos que haber optado por sistemas plurimetodológicos de investigación ha aportado consistencia al estudio, integrando los distintos niveles de análisis: declarativos, reflexivo y proposicional.

REFERENCIAS BIBLIOGRÁFICAS

- Abell, S. K., Appleton, K. y Hanuscin, D. (2010). *Designing and teaching the elementary science methods course*. Nueva York: Teaching and Learning in Science Series.
- Albert, M. J. (2007). *La investigación educativa. Claves teóricas*. Madrid: McGraw-Hill.
- Bardin, L. (1996). *El análisis de contenido*, 2.ª ed. Madrid: Akal.
- Beavers, A. S., Lounsbury, J. W., Richards, J. K., Huck, S. W., Skolits, G. J. y Esquivel, S. L. (2013). Practical considerations for using exploratory factor analysis in educational research. *Practical Assessment, Research y Evaluation*, 18(6), 1-13. <http://pareonline.net/getvn.asp?v=18&n=6>
- Bennett, R. (2011). Formative assessment: a critical review. *Assessment in Education: Principles, Policy y Practice*, 18(1), 5-25.
<https://doi.org/10.1080/0969594x.2010.513678>.
- Bisquerra, R. (2009). *Metodología de la investigación educativa*. Madrid: La Muralla.
- Black, P. y Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31.
<https://doi.org/10.1007/s11092-008-9068-5>.
- Buck, G. A., Trauth-Nare, A. y Kaftan, J. (2010). Making formative assessment discernable to pre-service teachers of science. *Journal of Research in Science Teaching*, 47(4), 402-421.
<https://doi.org/10.1002/tea.20344>.
- Colás, M. P., Buendía, L. y Hernández, F. (2009). *Competencias científicas para la realización de una tesis doctoral. Guía metodológica de elaboración y presentación*. Barcelona: Davinci.
- COSCE, Confederación de Sociedades Científicas de España (2011). *Informe ENCIENDE: Enseñanza de las Ciencias en la Didáctica Escolar para edades tempranas en España*. http://www.cosce.org/pdf/Informe_ENCIENDE.pdf
- Creswell, J. W. y Plano Clark, V. L. (2010). *Designing and Conducting Mixed Methods Research*, 2.ª ed. Thousand Oaks: Sage Publications.
- Dustchl, R., Maeng, S. y Sezen, A. (2011). Learning progressions and teaching sequences: review and analysis. *Studies in Science Education*, 47(2), 123-182.
<https://doi.org/10.1080/03057267.2011.604476>
- Escudero, T. (2016). La investigación evaluativa en el siglo xxi: un instrumento para el desarrollo educativo y social cada vez más relevante. *RELIEVE*, 22(1), art. 4, 1-21.
<http://dx.doi.org/10.7203/relieve.22.1.8164>
- Ezquerro, A. y Rodríguez-Marín, F. (2013). Aprender a enseñar ciencias a través del uso del video en formación inicial. *Investigación en la escuela*, 80, 67-76.
- Friedrichsen, P. J., van Driel, J. H. y Abell, S. K. (2011). Taking a closer look at science teaching orientations. *Science Education*, 95(2), 358-376.
<https://doi.org/10.1002/sce.20428>

- García, F. F. (2000). Los modelos didácticos como instrumento de análisis y de intervención en la realidad educativa. Biblio 3W. *Revista Bibliográfica de Geografía y Ciencias Sociales*, V(207). <http://www.ub.edu/geocrit/b3w-207.htm>
- Goodrum, D., Hackling, M. y Rennie, L. (2001). *The status and quality of teaching and learning of science in Australian schools*. Canberra: Department of Education, Training and Youth Affairs.
- Hamed, S. (2013). ¿Qué ideas tienen los futuros maestros de primaria acerca de qué y cómo enseñar y evaluar en ciencias? En *Actas IX Congreso Internacional sobre investigación en didáctica de las ciencias* (pp. 1726-1730). Girona.
- Hargreaves, E. (2005). Assessment for learning? Thinking outside the (black) box. *Cambridge Journal of Education*, 35(2), 213-224.
<https://doi.org/10.1080/03057640500146880>
- Harlen, W. (2013). *Assessment and Inquiry-Based Science Education: Issues in Policy and Practice*. Trieste, Italy: Global Network of Science Academies. Obtenido de <http://www.lulu.com/content/paperback-book/assessment-inquiry-based-science-education-issues-in-policy-and-practice/13672365>
- Hattie, J. (2009). *Visible learning: A synthesis of meta-analyses in education*. Londres: Routledge.
- López-Lozano, L. (2017). *Estudio de la evolución del conocimiento didáctico de estudiantes del Grado de Educación Primaria sobre la evaluación en Ciencias* (tesis doctoral). Universidad de Sevilla.
- López-Lozano, L., Solís, E. y Azcárate, P. (2018). Evolution of ideas about assessment in science: Incidence of a formative process. *Research in Science Education*.
<https://doi.org/10.1007/s11165-016-9591-1>
- López-Lozano, L., Solís, E. y Fernández-Arroyo, J. (2018). Los cambios en las concepciones sobre evaluación en ciencias de futuros maestros de Primaria. *Revista Complutense de Educación*, 29(3), 847-864.
<https://doi.org/10.5209/rced.54061>
- Maclellan, E. (2004). Initial knowledge states about assessment: novice teachers' conceptualisations. *Teaching and Teacher Education*, 20(5), 523-535.
<https://doi.org/10.1016/j.tate.2004.04.008>
- Martín del Pozo, R., Porlán, R. y Rivero, A. (2011). The progression of prospective teachers' conceptions of school science content. *Journal of Science Teacher Education*, 22(4), 291-312.
<https://doi.org/10.1007/s10972-011-9233-4>
- Martín del Pozo, R., Rivero, R. y Azcárate, P. (2014). Las concepciones de los futuros maestros sobre la naturaleza, cambio y utilización didáctica de las ideas de los alumnos. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 11(3), 348-363.
https://doi.org/10.25267/rev_eureka_ensen_divulg_cienc.2014.v11.i3.06
- Morales, P. (2007). *Estadística aplicada a las ciencias sociales. La fiabilidad de los tests y escalas*. Madrid: Universidad Pontificia Comillas. <https://matcris5.files.wordpress.com/2014/04/fiabilidad-tests-y-escalas-morales-2007.pdf>
- Morales, P. (2012). *El tamaño del efecto (effect size): análisis complementarios al contraste de medias*. Madrid: Universidad Pontificia Comillas. <http://web.upcomillas.es/personal/peter/investigacion/Tama%fl0DelEfecto.pdf>
- NGSS Lead States (2013). *Next generation science standards: For states, by states*. Washington, DC: National Academies Press.
- Osses, S., Sánchez, I. e Ibáñez, F. (2006). Investigación Cualitativa en Educación: Hacia la Generación de Teoría a través del proceso analítico. *Estudios Pedagógicos XXXII*, 1, 119-133. <http://mingaonline.uach.cl/pdf/estped/v32n1/art07.pdf>
<https://doi.org/10.4067/s0718-07052006000100007>

- Porlán, R., Martín del Pozo, R., Rivero, A., Harres, J., Azcárate, P. y Pizzato, M. (2010). El cambio del profesorado de ciencias I: Marco teórico y formativo. *Enseñanza de las Ciencias*, 28(1), 31-46.
- Porlán, R. y Rivero, A. (1998). *El conocimiento de los profesores*. Sevilla: Diada.
- Remesal, A. (2011). Primary and secondary teachers' conceptions of assessment: A qualitative study. *Teaching and Teacher Education*, 27(2), 472-482.
<https://doi.org/10.1016/j.tate.2010.09.017>
- Rivero, A., Azcárate, P., Porlán, R., Martín del Pozo, R. y Harres, J. (2011). The Progression of Prospective Primary Teachers' Conceptions of the Methodology of Teaching. *Research in Science Education*, 41(5), 739-769.
<https://doi.org/10.1007/s11165-010-9188-z>
- Rivero, A., Martín del Pozo, R., Solís, E., Azcárate, P. y Porlán, R. (2017). Cambio del conocimiento sobre la enseñanza de las ciencias de futuros maestros. *Enseñanza de las Ciencias*, 35(1), 29-52.
<https://doi.org/10.5565/rev/ensciencias.2068>
- Rivero, A., Martín del Pozo, R., Solís, E. y Porlán, R. (2017). *Didáctica de las Ciencias Experimentales en educación primaria*. Síntesis: Madrid.
- Rivero, A., Porlán, R., Solís, E., Rodríguez-Marín, F., Hamed, S., Martín del Pozo, R., Ezquerra, A. y Azcárate, P. (2012). *Aprender a enseñar ciencias en primaria. Actividades de formación inicial de maestros para aprender a enseñar ciencias por investigación escolar*. Sevilla: Copiarte.
- Rocard, M., Csermely, P., Jorde, D., Lenzen, D., Walberg-Henriksson, H. y Hemmo, V. (2007). *Science Education Now: A Renewed Pedagogy for the Future of Europe*. Bruselas. http://ec.europa.eu/research/science-society/document_library/pdf_06/report-rocard-on-science-education_en.pdf
- Rodríguez-Marín, F., Fernández-Arroyo, J. y García-Díaz, J. E. (2014). Las hipótesis de transición como herramienta didáctica para la Educación Ambiental. *Enseñanza de las Ciencias*, 32(3), 303-318.
<https://doi.org/10.5565/rev/ensciencias.1137>
- Ruiz-Olabuénaga, J. I. (2012). *Metodología de la investigación cualitativa*, 5.ª ed. Bilbao: Publicaciones de la Universidad de Deusto.
- Sánchez-Gómez, M. C. (2015). La dicotomía cualitativo-cuantitativo: posibilidades de integración y diseños mixtos. *Campo Abierto*, vol. monográfico, 11-30.
- Sanmartí, N. (2012). *Evaluar para aprender. 10 Ideas claves*. Barcelona: Editorial Graò.
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14.
<https://doi.org/10.1002/9780470690048.ch10>
- Sickel, A. J. (2017). The 5e model as a framework for facilitating multiple teacher Education outcomes. En A. J. Sickel y S. B. Witzig (Eds.), *Designing and Teaching the Secondary Science Methods Course. An International Perspective* (pp. 11-31). Róterdam: Sense Publishers.
- Solís, E. y Porlán, R. (2003). Las concepciones del profesorado de ciencias de secundaria en formación inicial: ¿obstáculos o punto de partida? *Investigación en la Escuela*, 49, 5-22.
- Tashakkori, A. y Creswell, J. W. (2007). Editorial: The New Era of Mixed Methods. *Journal of Mixed Methods Research*, 1(1), 3-7.
<https://doi.org/10.1177/2345678906293042>
- Tójar, J. C. (2006). *Investigación cualitativa: comprender y actuar*. Madrid: La Muralla.
- Wang, J.-R., Kao, H.-L. y Lin, S.-W. (2010). Pre-service teachers' initial conceptions about assessment of science learning: The coherence with their views of learning science. *Teaching and Teacher Education*, 26(3), 522-529.
<https://doi.org/10.1016/j.tate.2009.06.014>

A piece of research on the evolution of teachers' didactic knowledge about assessment in science

Lidia López-Lozano

Escuela Universitaria de Osuna, Universidad de Sevilla, Sevilla, España.

lidialopez@us.es

Emilio Solís Ramírez

Departamento de Didáctica de las Ciencias Experimentales y Sociales, Universidad de Sevilla, Sevilla, España.

esolis@us.es

We present a work analyzing the change in didactic knowledge about assessment in science among pre-service teachers. This contribution is part of the results obtained with an I+D+i project. The purpose of this project was to develop and evaluate a teacher training model based on the research carried out about relevant curricular problems and on the interaction with innovative teaching practices from a constructivist perspective of the teaching-learning process.

During the twentieth century, the assessment process in the educational field has evolved (Escudero, 2016; Shepard, 2000) towards socio-constructivist learning theories. In this context, assessment is presented as a critical component of the process of teaching and learning as based on inquiry (Harlen, 2013). When talking about assessment, one traditionally thinks in terms of «measuring» the performance of students by focusing on the acquisition of conceptual contents. This might as well be the most «simplistic» view. Nevertheless, in the assessment process, we think that what really matters are its implicit purposes: one is «formative», helping or promoting the learning capacity, while the other is «summative» to take into account what was learned at a given time (Bennett, 2011).

The research methodology is mixed, with a concurrent design of qualitative and quantitative techniques by means of a triangulation process. We agree with Albert (2007), who proposes an empirical-analytical perspective, an interpretative conception to understand human behaviour and thus be able to interpret and understand educational phenomena rather than provide causal explanations. We must also keep in mind the critical current that tries to reveal beliefs, values and assumptions at the basis of educational practice. In our study, the existence of elements and assumptions which are acceptable in both cases leads us to propose an appropriate combination of the two, i. e. a multimethod approach (Creswell and Plano, 2010; Tashakkori and Creswell, 2007). The procedure followed is valuable and useful for documenting the evolution in teachers' conceptions. The study focuses on 347 students taking the degree in Primary Education at the University of Seville and who study the subject of Didactic of Experimental Sciences. We analyse questionnaires (pre and post-test), sets of open-answer reflective questions and the teaching proposals prepared by those students.

In relation to the changes produced and their degree of coherence, we can observe that in the beginning, there is some asymmetry between what they declare when completing the questionnaire (at the identification level) and what they propose in their teaching proposals (at the declarative-proposal level). Furthermore, there are discrepancies between their teaching proposals and what they state in the set of reflective questions when they analyse or reflect on their approaches (at the declarative-reflexive level). They show agreement with a type of assessment that favours learning, but they design it closer to their references as students (Buck et al., 2010; Wang et al., 2010). The results at the end of the training show changes related to a greater coherence between the declarative and factual knowledge of the teachers. These changes are stimulated to a large extent by the reflection on their design.