



Universitat Autònoma de Barcelona

ADVERTIMENT. L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons:  http://cat.creativecommons.org/?page_id=184

ADVERTENCIA. El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons:  <http://es.creativecommons.org/blog/licencias/>

WARNING. The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license:  <https://creativecommons.org/licenses/?lang=en>

Tesis doctoral

Evaluación e Identidad Docente: impacto de una formación sobre evaluación auténtica de competencias

Autora: Esperanza Mejías Macías

Director: Carles Monereo Font

**Programa de Doctorado Interuniversitario en Psicología de la
Educación**

Departamento de Psicología Básica, Evolutiva y de la Educación

Facultad de Psicología

2019



**Universitat Autònoma
de Barcelona**

Agradecimientos

A todas las personas que conozco que creen que un mundo mejor es posible gracias a la educación. A aquellas que tanto me han enseñado sobre la enseñanza, la vida, el amor y la ética. Y, especialmente, a las que me han acompañado a aprender a mirar el mundo con ojos de científica.

Resumen

Esta investigación tiene como propósito ayudar a mejorar la competencia en evaluación de los profesores a través de una formación sobre evaluación auténtica que impacte en su identidad docente, específicamente, en su posicionamiento sobre evaluación, como medida de sostenibilidad del cambio. En un primer estudio comprobamos la eficacia de un instrumento formativo que ayuda a los docentes a transformar sus exámenes tradicionales en pruebas de evaluación auténticas. Implementamos una formación en la que participaron 206 docentes de distintos centros educativos de enseñanza obligatoria a lo largo de dos cursos escolares. Evaluamos las pruebas diseñadas durante la formación y, a partir de un análisis correlacional, se compararon las puntuaciones obtenidas en las distintas versiones con variables independientes como la experiencia docente o la materia. Los resultados indican un grado significativo de mejora en las pruebas de evaluación diseñadas en sus características de evaluación de competencias, evaluación formativa y atención a la diversidad. En un segundo estudio identificamos y caracterizamos el posicionamiento sobre evaluación según se oriente a una evaluación tradicional o a una evaluación auténtica, para explorar su relación con la permanencia en el uso y aplicación de las pruebas auténticas diseñadas en la formación. Aplicamos un cuestionario a una muestra de 20 docentes de entre todos los que participaron en la formación. Los resultados muestran que los docentes, de forma mayoritaria, sostienen posiciones con una tendencia a la evaluación auténtica y que sus pruebas actuales mantienen características de evaluación de competencias, evaluación formativa y atención a la diversidad.

Sumario

ÍNDICE TABLAS.....	8
ÍNDICE DE FIGURAS	10
INTRODUCCIÓN	11
CAPÍTULO 1. FUNDAMENTACIÓN TEÓRICA	16
1.1. LA EVALUACIÓN DE Y EN LOS PROCESOS DE ENSEÑANZA Y APRENDIZAJE	16
1.1.1. <i>De la evaluación tradicional a la evaluación alternativa.....</i>	<i>18</i>
1.1.1.1. Tipos de evaluación.....	18
1.1.1.2. Desde la evaluación alternativa.....	23
1.1.2. <i>Los componentes de la evaluación.....</i>	<i>27</i>
1.1.2.1. Para qué evaluamos.....	29
1.1.2.2. Qué evaluamos.	32
1.1.2.3. En referencia a qué evaluamos.....	36
1.1.2.4. Con qué evaluamos.....	38
1.1.2.5. Cuándo evaluamos.....	43
1.1.2.6. Quiénes evaluamos.....	46
1.1.3. <i>El proceso de evaluación.....</i>	<i>49</i>
1.1.3.1. La retroalimentación durante la evaluación.....	54
1.2. LA EVALUACIÓN AUTÉNTICA	57
1.2.1. <i>Bases psicoeducativas, principios y características.....</i>	<i>59</i>
1.2.1.1. La autenticidad requiere evaluar competencialmente.....	64
1.2.1.2. La autenticidad requiere que la evaluación sea formativa.....	67
1.2.1.3. La autenticidad requiere evaluar de forma inclusiva.....	70
1.2.2. <i>Experiencias de evaluación auténtica.....</i>	<i>72</i>
1.2.3. <i>La evaluación auténtica en el aula.....</i>	<i>76</i>
1.3. LA COMPETENCIA EN EVALUACIÓN DE LOS DOCENTES	80
1.3.1. <i>La importancia de una alfabetización en evaluación.....</i>	<i>81</i>
1.3.2. <i>La necesidad de ser competentes en evaluación.....</i>	<i>85</i>
1.4. LA FORMACIÓN CONTINUA SOBRE EVALUACIÓN	90
1.4.1. <i>La formación permanente y su transferencia.....</i>	<i>91</i>
1.4.2. <i>La formación sobre evaluación.....</i>	<i>95</i>
1.4.2.1. Maestros que desarrollan evaluaciones para el aprendizaje: impacto en el logro estudiantil.....	97
1.4.2.2. Entendiendo cómo los profesores se involucran en la evaluación formativa.....	98
1.4.2.3. Evaluación formativa y compartida en educación: experiencias de éxito en todas las etapas educativas.....	99
1.4.2.4. Evaluar para aprender: la evaluación al servicio del aprendizaje del alumno.....	99
1.4.3. <i>Formación docente sobre evaluación: Proyecto GAPPISA.....</i>	<i>101</i>

1.4.3.1. Origen del proyecto.	102
1.4.3.2. ¿En qué se basa el Proyecto GAPPISA?.....	103
1.4.3.3. ¿En qué consiste el Proyecto GAPPISA?	108
1.5. LA IDENTIDAD DOCENTE Y EL POSICIONAMIENTO SOBRE EVALUACIÓN.....	124
1.5.1. <i>La Identidad Docente.</i>	125
1.5.2. <i>El posicionamiento docente sobre evaluación.</i>	127
1.5.3. <i>Dispositivos para la reconstrucción del posicionamiento en evaluación</i>	138
CAPÍTULO 2. ESTUDIO 1: EVALUANDO LAS EVALUACIONES. EFICACIA DE UNA GUÍA PARA LA ELABORACIÓN DE PRUEBAS AUTÉNTICAS	144
2.1. INTRODUCCIÓN Y OBJETIVOS.....	144
2.2. MÉTODO DEL ESTUDIO.....	150
2.2.1. <i>Contexto.</i>	150
2.2.2. <i>Participantes.</i>	151
2.2.3. <i>Procedimiento.</i>	151
2.2.4. <i>Instrumentos.</i>	152
2.2.5. <i>Análisis de datos.</i>	153
2.3. RESULTADOS	154
2.3.1. <i>Análisis de las mejoras en las pruebas de evaluación.</i>	155
2.3.1.1. Cambios en las pruebas.....	155
2.3.1.2. Relación con otras variables.	158
2.3.2. <i>Descripción de las mejoras en las pruebas de evaluación.</i>	159
2.3.2.1. Las mejoras a nivel de competencialidad.....	160
2.3.2.2. Las mejoras a nivel de poder formativo.....	162
2.3.2.3. Las mejoras a nivel de inclusividad.	164
2.3.3. <i>Las pseudomejoras en las pruebas de evaluación.</i>	166
2.4. DISCUSIÓN DE LOS RESULTADOS	169
2.5. CONCLUSIONES, LIMITACIONES Y PROSPECTIVA	176
CAPÍTULO 3. ESTUDIO 2: CAMBIOS EN EL POSICIONAMIENTO DOCENTE SOBRE EVALUACIÓN A PARTIR DE UNA FORMACIÓN SOBRE EVALUACIÓN AUTÉNTICA.....	179
3.1. INTRODUCCIÓN Y OBJETIVOS.....	179
3.1.1. <i>Dispositivos para el cambio de posicionamiento en evaluación</i>	183
3.2. MÉTODO DEL ESTUDIO.....	187
3.2.1. <i>Contexto y procedimiento.</i>	187
3.2.2. <i>Participantes.</i>	188
3.2.3. <i>Instrumentos de recogida y análisis de datos.</i>	188
3.3. RESULTADOS	190
3.3.1. <i>Posicionamiento de los docentes sobre evaluación</i>	191
3.3.2. <i>Incidentes sobre evaluación.</i>	197

3.3.3. <i>Permanencia en el diseño y aplicación de la prueba auténtica y competencial.</i>	199
3.4. DISCUSIÓN DE LOS RESULTADOS.	201
3.5. CONCLUSIONES, LIMITACIONES Y PROSPECTIVA	208
CAPÍTULO 4. CONCLUSIONES	211
REFERENCIAS	219
ANEXOS	245
ANEXO 1: GUÍA GAPPISA.....	246
ANEXO 2. RÚBRICA DE EVALUACIÓN DE LAS PRUEBAS AUTENTIFICADAS	250
ANEXO 3. INFORME FINAL SOBRE LA APLICACIÓN DE LA PRUEBA	254
ANEXO 4. ANÁLISIS DE LOS ESTUDIOS SOBRE CONCEPCIONES DE LOS DOCENTES SOBRE EVALUACIÓN.....	259
ANEXO 5. CUESTIONARIO "MI POSICIONAMIENTO SOBRE EVALUACIÓN..."	264
ANEXO 6. INSTRUMENTO DE ANÁLISIS DEL POSICIONAMIENTO DOCENTE SOBRE EVALUACIÓN	270

Índice tablas

Tabla 1: Tipos de evaluación según los componentes que la caracterizan	19
Tabla 2: Características de la evaluación tradicional y de la evaluación alternativa	24
Tabla 3: Cambios que debe asumir la evaluación para responder a las nuevas teorías sobre aprendizaje	25
Tabla 4: Componentes que definen la evaluación y preguntas conductoras para definirlos	28
Tabla 5: Inventario de instrumentos de evaluación según los objetivos de evaluación	39
Tabla 6: Ciclo de evaluación de aprendizajes en el aula	52
Tabla 7: Dimensiones y características de la evaluación auténtica según Frey et al. (2012)	62
Tabla 8: Dimensiones y características de la evaluación auténtica según Villarroel et al. (2017)	63
Tabla 9: Descripción de las características de una evaluación auténtica que evalúa competencias	65
Tabla 10: Descripción de las características de una evaluación auténtica orientada al aprendizaje	68
Tabla 11: Descripción de las características de una evaluación auténtica que atiende la diversidad	70
Tabla 12: Comparación entre el sistema de evaluación tradicional y el de evaluación auténtica	76
Tabla 13: Características de las pruebas PISA según año y competencia clave evaluada	105
Tabla 14: Relación entre las características para una formación docente eficiente y las características del Proyecto GAPPISA	108
Tabla 15: Programación del Proyecto GAPPISA por fases y actividades	111
Tabla 16: Descripción de los recursos y los instrumentos formativos del Proyecto GAPPISA	115
Tabla 17: Comparación entre las características de la Guía GAPPISA y las pruebas PISA	119
Tabla 18: Descripción de las características de la Guía GAPPISA	120
Tabla 19: Representaciones que conforman las concepciones sobre evaluación clasificadas según el componente al que hagan referencia	132
Tabla 20: Relación entre las dimensiones de la identidad docente y los componentes de la evaluación que pueden asociarse a cada una de ellas	134
Tabla 21: Continuo de variables que definen las categorías del posicionamiento docente sobre evaluación	136
Tabla 22: Relación entre las dimensiones, las variables y los indicadores de análisis del estudio	153
Tabla 23: Número de entregas de los docentes según la versión de la prueba	155

Tabla 24: Datos estadísticos de la puntuación final de todas las pruebas en cada una de sus versiones	156
Tabla 25: Relación entre las categorías, las variables y los indicadores de análisis del estudio	189
Tabla 26: Posicionamiento sobre evaluación de los docentes según su tendencia	192
Tabla 27: Tipos de incidentes críticos con relación a la evaluación.....	198

Índice de Figuras

Figura 1. Sistema de evaluación de competencias según Jornet, González, Suárez y Peraless (2011).....	51
Figura 2. Esquema sobre las fases del Proyecto GAPPISA	111
Figura 3. Nivel alcanzado en cada una de las pruebas según las categorías de competencialidad, poder formativo e inclusividad.....	158
Figura 4. Prueba competencial de matemáticas, de 2º ESO, después de la formación.	161
Figura 5. Fragmentos de la prueba competencial de lengua y literatura, 3º ESO, después de la formación.	163
Figura 6. Fragmentos de la prueba competencial de historia, 4º ESO, después de la formación.....	165
Figura 7. Posicionamiento sobre evaluación de los docentes por dimensión.....	192
Figura 8. Tipo de evaluación según las categorías que conforman las dimensiones del posicionamiento docente sobre evaluación.	193
Figura 9. Tendencia en la variable finalidad de la evaluación.....	194
Figura 10. Tendencia en la variable objeto de la evaluación.	194
Figura 11. Tendencia en la variable agente que evalúa	195
Figura 12. Tendencia en la variable instrumentos de evaluación	195
Figura 13. Tendencia en la variable sentimientos respecto a la evaluación	196
Figura 14. Tendencia en la variable preocupaciones sobre evaluación	197
Figura 15. Permanencia sobre el diseño y la aplicación de pruebas que cumplen las características de la Guía GAPPISA	200
Figura 16. Características que cumplen las pruebas y actividades de evaluación de los profesores según el modelo GAPPISA	201

Introducción

Las funciones de los profesionales de la Psicología de la Educación y de la Pedagogía son muchas y variadas, pero todas comparten un propósito común: acompañar a la comunidad educativa en su proceso actualización constante para la calidad del sistema educativo. De este objetivo se desprenden muchas líneas de actuación según el nivel de intervención en el que nos situemos y el contexto en el que intervengamos. En esta investigación, ponemos el foco en los procesos de enseñanza y aprendizaje y en la actualización pedagógica de los maestros y profesores encargados de estos. Como metodología principal, optamos por la formación y asesoramiento a maestros y profesores en activo que están dispuestos a mejorar sus prácticas para linearlas con el enfoque de enseñanza competencial y escuela inclusiva.

Son muchos y variados los cambios que se deben introducir para conseguir una metodología de enseñanza basada en el desarrollo de competencias que, además, dé respuesta a los perfiles de aprendizaje de los alumnos. Los hay más específicos, como la introducción de metodologías activas y más generales como el diseño de programas de atención a la diversidad, por tanto, algunos tienen mayor incidencia que otros en la realidad de las aulas. En este sentido, encontramos en la introducción de cambios en la evaluación una innovación que nos permite repensar la enseñanza y mejorar los procesos de aprendizaje de los alumnos. Al respecto, varios autores señalan que transformar la evaluación tiene efectos retroactivos en la enseñanza, en tanto que son las caras de una misma moneda, y que una evaluación formativa e integrada en los procesos de enseñanza y aprendizaje ayuda a los alumnos a aprender a autorregularse y ser más competentes (López-Pastor y Pérez-Pueyo, 2017; Monereo, 2009; Sanmartí, 2010a, 2010b; Swaffield, 2011).

Así pues, en esta investigación partimos del marco de la evaluación alternativa conceptualizada como el conjunto de procesos y prácticas de evaluación centradas en recoger evidencias, analizarlas y emitir juicios que ayuden a los docentes a tomar decisiones sobre cómo mejorar su enseñanza y a los alumnos sobre cómo mejorar su aprendizaje, participando estos también del proceso (Álvarez, 2001;

Mateo y Martínez, 2008). Dentro del paraguas de la evaluación alternativa encontramos diversidad de prácticas que responden a este objetivo, aunque no todas ayudan de igual manera al desarrollo de las competencias.

Por ello, vamos un paso más allá y optamos por la evaluación auténtica, una evaluación basada en el desempeño y orientada al aprendizaje, que da respuesta a las necesidades de los nuevos enfoques de enseñanza y las teorías actuales sobre el aprendizaje. Sus características principales identificadas en la literatura especializada en este modelo, como son el realismo, la complejidad cognitiva, los criterios compartidos y la participación del alumnado, demuestran que es un modelo válido y ajustado a los retos de la evaluación de competencias (Ashford-Rowea, Herrington y Brown, 2014; Brown, 2015; Darling-Hammond y Snyder, 2000; Gulikers, Bastiaens y Kirschner, 2004; Frey, Schmitt y Allen, 2012; A. Herrington y J. Herrington, 2006; Villarroel, Bloxham, D.Bruna, C.Bruna y Herrera-Seda, 2017).

Todo ello nos conduce al requerimiento de que los docentes implementen y gestionar estratégicamente prácticas de evaluación auténtica para alinear la evaluación con los nuevos enfoques competenciales. Sin embargo, la realidad es que maestros y profesores todavía distan mucho de saber cómo integrar la evaluación en sus programaciones habituales, debido a que, tanto los planes de formación inicial del profesorado como las experiencias que han vivido como estudiantes, se corresponden, mayoritariamente, con una evaluación tradicional (Clark y Rust, 2006; Deneen y Brown, 2016; DeLuca, LaPointe-McEwan y Luhanga, 2016). De esta realidad, emerge la necesidad de acompañar a los docentes en activo en el desarrollo de su competencia en evaluación, para suplir la falta de alfabetización en evaluación y las dificultades para transformara sus prácticas (Mellati y Khademi, 2018; Tonda y Medina, 2013; Sanahuja y Sánchez-Tarazaga, 2018)

Con ello, llegamos a nuestro primer objetivo que consiste en formar a maestros y profesores en evaluación auténtica para que puedan evaluar competencias de manera formativa e inclusiva, como explicaremos a continuación. Ahora bien, para alcanzar este objetivo, no es suficiente con aplicar nuevas prácticas, los docentes deben desarrollar su competencia en evaluación, es decir, las habilidades y el

conocimiento estratégico suficientes como para sostener estas innovaciones en el tiempo, superando las resistencias y los contratiempos que puedan surgir. En este sentido, se ha demostrado que cambiar la estrategia no es suficiente y que, para que los docentes puedan persistir en su objetivo de forma competente, es necesario un cambio en la identidad docente (Monereo, Badia, Bilbao, Cerrato y Weise, 2009).

Entendemos la identidad como el conjunto de representaciones que tiene el profesional sobre "ser docente" de formar relativamente estable y la construcción de posiciones durante la relación con el contexto (Badia, García y Meneses, 2017; Hermans, 2009; Monereo y Badia, 2011). Dentro de este marco, identificamos las posiciones sobre evaluación como aquellas situaciones específicas en las que el docente actúa según las concepciones, las estrategias y los sentimientos que tenga relativos a la evaluación. Y esto es importante, porque la literatura ha demostrado que un cambio profundo en este posicionamiento, es decir, la adopción de una nueva posición sobre evaluación puede ayudar al docente a superar los incidentes críticos por los que pase (Nail, Gajardo y Muñoz, 2012; Monereo, Weise y Álvarez, 2013) y, por tanto, a sostener en el tiempo la innovación educativa.

Con ello llegamos a nuestro segundo objetivo consistente en identificar y describir cuáles son los elementos constituyentes del posicionamiento sobre la evaluación de los docentes. El conocimiento de estas posiciones evaluativas nos permitirá demostrar si un cambio en la identidad se corresponde con una sostenibilidad de innovaciones introducidas en los procesos de enseñanza y aprendizaje. De esta manera, podremos avanzar en la investigación sobre cómo diseñar formaciones que impacten en la identidad docente, proponiendo el estudio de las posiciones docentes como estrategia efectiva de transformación de la realidad educativa y del desarrollo de docentes competentes.

Para comprobar estas hipótesis y dar respuesta a los objetivos planteados diseñamos e implementamos una formación sobre evaluación auténtica, denominada Proyecto GAPPISA en coincidencia con el acrónimo que da nombre al instrumento principal, *Guía para el Análisis de Pruebas desde la perspectiva PISA* -a partir de ahora guía GAPPISA- (Álvarez-Bernárdez y Monereo, 2016; Duran y Monereo, 2008; Mejías y Monereo, 2018; Monereo y Morcillo, 2011). En esta

formación los docentes, mediante el acompañamiento de un equipo de asesores, transforma un examen en una prueba de evaluación auténtica, la lleva a la práctica en un contexto real de aula y reflexiona sobre los resultados conseguidos y los cambios que ha habido en los aprendizajes de los alumnos y en su forma de enseñar.

En un primer estudio, analizamos cuáles han sido los resultados obtenidos en una muestra de 206 docentes, describiendo los cambios más significativos que han realizado en sus prácticas de evaluación y examinando la incidencia de la formación. En un segundo estudio, seleccionamos una muestra de los docentes participantes en la formación con el objetivo de identificar y describir cuáles son las posiciones sobre evaluación y los incidentes críticos que más les afectan. De esta manera, queremos comprobar la permanencia de los cambios en las prácticas de evaluación y cómo estos han influido en la construcción de la posición evaluativa de los docentes.

Para dar cuenta de esta investigación, en los siguientes capítulos describimos el marco teórico que permite justificar y construir esta investigación y comunicamos los resultados de los dos estudios, sus limitaciones y las nuevas líneas de investigación que emergen de la misma. En el primer capítulo, la fundamentación teórica, nos situaremos conceptualmente en la evaluación en y de los procesos de enseñanza y aprendizaje, definiendo cuáles son los términos más comunes, y describiendo los componentes y el proceso necesario para hacer posible una evaluación orientada al aprendizaje. En segundo lugar, ahondaremos en la evaluación auténtica, describiendo sus bases y características en profundidad. En tercer lugar, indicaremos cuáles son los condicionantes que han llevado a los docentes a necesitar mejorar su alfabetización en evaluación y justificaremos la necesidad de desarrollar la competencia en evaluación de los docentes. En cuarto lugar, hablaremos de la formación continua de los docentes y cómo conseguir una transferencia de los aprendizajes realizados en esta a la práctica; analizaremos formaciones sobre evaluación y describiremos en profundidad la formación sobre evaluación auténtica que implementamos. En quinto lugar, expondremos el concepto de identidad docente y examinaremos en qué consiste el posicionamiento docente sobre evaluación, para finalizar el apartado indicando cuáles son los

dispositivos de cambio que favorecen una transformación de la posición evaluadora.

Seguidamente, en los capítulos segundo y tercero, desarrollaremos el marco empírico en dos estudios conectados entre sí. Un primer estudio titulado *Evaluando las evaluaciones: eficacia de una guía para la elaboración de pruebas auténticas*. Y un segundo estudio titulado *Cambios en el posicionamiento docente sobre evaluación a partir de una formación sobre evaluación auténtica*. Ambos estudios cuentan con una justificación teórica, un marco metodológico, una discusión de los resultados analizados, y unas conclusiones y las limitaciones y prospectiva finales.

Finalmente, exponemos unas conclusiones generales de la investigación en las que explicaremos de qué manera hemos alcanzado los objetivos, las aportaciones e implicaciones que los resultados tienen para la práctica, las limitaciones de la investigación y las futuras líneas de investigación que emergen a partir de los resultados obtenidos.

Capítulo 1. Fundamentación Teórica

1.1. La Evaluación de y en los Procesos de Enseñanza y Aprendizaje

Evaluar es una acción que realizamos de forma cotidiana que implica enjuiciar algo. En el momento en que debemos tomar una decisión, por muy sencilla que esta sea, normalmente valoramos las opciones disponibles, clasificamos, medimos, juzgamos, suponemos, ponderamos, etc. para poder decir con cierta seguridad: esto es lo que me conviene, es decir, para tomar una decisión. A veces, hacemos esto de forma muy intuitiva, siguiendo patrones de actuación que creemos que nos han funcionado en otras ocasiones y de los cuáles no tenemos mucha consciencia. Otras veces, necesitamos establecer relaciones más complejas y sistemáticas para prever con más precisión qué conseguiremos y de qué forma queremos hacerlo, según valoremos cómo es de importante la decisión que tenemos que tomar. Pero este tipo de evaluación se realiza de forma intuitiva y mayoritariamente solitaria.

Cuando hablamos de evaluación en un contexto educativo, sí hablamos necesariamente de una evaluación intencional que permite a los participantes tomar decisiones para mejorar. Además, está formada por un conjunto de acciones y de decisiones tomadas con el fin de juzgar y mejorar la calidad educativa llevadas a cabo de forma planificada. Según Lukas y Santiago (2009):

La evaluación es el proceso de identificación, recogida y análisis de información relevante – que podrá ser cuantitativa o cualitativa-, de manera sistemática, rigurosa, planificada, dirigida, objetiva, creíble, fiable y válida para emitir juicios de valor basados en criterios y referencias preestablecidos para determinar el valor y el mérito del objeto educativo en cuestión a fin de tomar decisiones que ayuden a optimizar el mencionado objeto. (pág. 91)

Como vemos, es necesario preguntarse el qué, el cuándo, el cómo, el quién y el por qué de la evaluación, planificarla y sistematizarla para que sea realmente una evaluación funcional. Ahora bien, esta conceptualización de la evaluación es aplicable a otros contextos, como el empresarial, donde es muy común que agencias especializadas evalúen de la calidad de la empresa para mejorar sus

resultados. Entonces, si queremos hablar de evaluación en el ámbito educativo, ¿cuál es el aspecto diferenciador?

Claramente, el objeto de la evaluación es un aspecto importante en esta diferenciación. Dentro del ámbito educativo, podemos evaluar al profesorado, el currículum, los proyectos educativos, a los alumnos e incluso el propio sistema. En este estudio nos vamos a centrar en la evaluación de los procesos de enseñanza y aprendizaje. un así, todavía hay un largo camino que recorrer para poder conocer cómo podemos conceptualizarla, considerando que tiene sus especificidades. Para esto, claramente, el propósito de la evaluación es el que marca la diferencia. No es lo mismo evaluar los procesos de enseñanza y aprendizaje para medirlos y certificar los resultados, que hacerlo para obtener evidencias que nos ayuden a saber de qué manera mejorarlos. Así pues, nos situaremos en la denominada evaluación formativa, entendida como aquella que sirve para mejorar el aprendizaje del alumnado, la competencia del docente y los propios procesos de enseñanza y aprendizaje (López-Pastor, 2012).

Desde esta perspectiva, una primera consideración para tener en cuenta es que todos los verbos que utilicemos para definir la evaluación. alejados a esta intencionalidad, deberían resituarnos. Nos indican, claramente, que estamos evaluando, pero no como una actividad dentro del proceso de enseñanza y aprendizaje. Siguiendo el discurso de Álvarez (2001), evaluar en una actividad de aprendizaje no puede suponer otra cosa que aportar información y ayudas para que el aprendiz mejore después de comparar sus resultados. De poco sirve para el que no sabe cómo hacerlo, saber que lo hizo muy bien o un poquito peor que la última vez, lo que realmente necesita es saber qué ha sucedido durante el aprendizaje que le ha llevado a esos resultados, para poder cambiarlo. Y podemos replicar exactamente la misma aseveración con la palabra enseñar y con los docentes como evaluadores de su propia enseñanza.

En resumen, en este apartado vamos a definir una evaluación educativa de y en los procesos de enseñanza y aprendizaje orientada a mejorar los aprendizajes de los alumnos. Para ello, definiremos cuáles son los modelos de evaluación y cómo construir un sistema de evaluación. En un primer punto, describiremos los tipos de

evaluación más comunes y, en un segundo apartado, concretaremos cuáles son los componentes y el proceso necesario para conseguirlo.

1.1.1. De la evaluación tradicional a la evaluación alternativa.

El término evaluación educativa es altamente polisémico dado que ha ido evolucionando en paralelo a los cambios de las leyes educativas, los nuevos descubrimientos sobre el aprendizaje y las nuevas metodologías de enseñanza. Así pues, son muchos los términos con los que la literatura se refiere a la evaluación: tradicional, alternativa, sumativa, formativa, formadora, de ejecución, auténtica, relativa, criterial, de evidencias, de proceso, etc. Sin embargo, esta diversidad, denominada en muchos estudios como tipos de evaluación, no es más que una caracterización de un mismo término.

Es decir, desde nuestra perspectiva, no es que existan muchas y diferentes evaluaciones, sino que existen muchas posibilidades para concretar y construir sistemas de evaluación de los procesos de enseñanza y aprendizaje. Y esta concreción hace referencia al enfoque utilizado, los componentes claves y el proceso que los concreta. En este punto vamos a hacer una revisión de los términos más utilizados, analizando qué aspectos los definen y cómo y por qué han evolucionado de un enfoque tradicional a uno denominado alternativo. Seguidamente, vamos a definir cuáles son los componentes de la evaluación y qué proceso debemos seguir si optamos por una evaluación orientada al aprendizaje.

1.1.1.1. Tipos de evaluación.

Como hemos avanzado, los términos que se utilizan para definir de evaluación son muchos y tienen su origen en los diferentes aspectos que la conforman. Según la finalidad, el momento en que se evalúe, el alcance de la evaluación, el agente, el enfoque o el estándar de comparación se utiliza un término u otro (ver Tabla 1). Esta polisemia responde claramente a una evolución sobre cómo se entiende la evaluación y a qué necesidades responde, durante la cuál se fueron incorporando nuevos términos que pudieran dar respuesta a los nuevos objetivos. Por lo tanto, se hace necesario, como primer paso, analizar cuáles han sido los términos que

han sido más representativos de este cambio. A continuación, vamos a definirlos y a explicar cuál ha sido la dirección que ha tomado esta evolución.

Tabla 1

Tipos de evaluación según los componentes que la caracterizan

Componentes	Tipos de evaluación
Finalidad	Evaluación Diagnóstica Evaluación Formativa Evaluación Formadora Evaluación Sumativa
Momento	Evaluación Inicial Evaluación Continua Evaluación Final Evaluación Integrada
Agente	Heteroevaluación Autoevaluación Coevaluación Evaluación Compartida
Actividad	Evaluación Observacional Evaluación Experimental Evaluación del Desempeño Evaluación basada en Evidencias
Referentes de valoración	Evaluación Normativa Evaluación Criterial
Extensión	Evaluación Global Evaluación Parcial
Enfoque metodológico	Evaluación Cualitativa Evaluación Cuantitativa

La noción de evaluación sumativa - *summative assessment* - describe muy bien la idea de medir, al finalizar las prácticas educativas, los aprendizajes alcanzados por los alumnos, generalmente cuantificándolos y graduándolos de menor a mayor grado de consecución, según unos criterios no compartidos o según la comparación de los resultados entre los alumnos. Por tanto, a esta tipología, normalmente van asociadas la evaluación final, la heteroevaluación, la evaluación cuantitativa y la evaluación normativa. A veces, cuando se realizan diversas evaluaciones de esta manera, durante una misma unidad didáctica, también se asocia con la evaluación continua. Esta interpretación se denomina en otras ocasiones y de manera errónea, evaluación formativa, puesto que parece valorar el proceso cuando, en realidad, se están evaluando distintos productos intermedios, modelo que podríamos considerar una evaluación sumativa por fases.

La gran diferencia entre la evaluación sumativa y la evaluación formativa es que el objetivo de esta última es analizar las evidencias de aprendizaje para mejorarlo, de forma cualitativa, compartida, criterial y continua, llegando a integrar la evaluación en los procesos de enseñanza y aprendizaje, y poniéndolos a la vez a su servicio. En palabras de López-Pastor y Pérez-Pueyo (2017), la evaluación formativa es:

Todo proceso de evaluación cuya finalidad principal es mejorar los procesos de enseñanza-aprendizaje que tienen lugar. Sirve para que el alumnado aprenda más (y/o corrija sus errores) y para que el profesorado aprenda a trabajar mejor (a perfeccionar su práctica docente). (p. 36)

Es decir, una evaluación que interpela tanto a estudiantes como profesores a mejorar; a los primeros, a optimizar su aprendizaje a través de la reflexión y autorregulación, y a los segundos, a precisar su rol a través del análisis de sus metodologías y propósitos (Black, 2015; Black y William, 2009). A partir de este enfoque de evaluación de y en los procesos de enseñanza y aprendizaje, han surgido un conjunto de tipologías que ponen su foco especialmente en mejorar el aprendizaje mediante la participación de los alumnos en la evaluación. Encontramos términos como la evaluación para el aprendizaje –*assessment for learning*–, la evaluación orientada al aprendizaje –*learning-oriented assessment*–, la evaluación centrada en el aprendizaje o la evaluación formadora. Se caracterizan por promover la reflexión y la regulación de los estudiantes y favorecer que se autoevalúen (Sanmartí, 2010a, 2010b; Santos-Guerra, 2003; Swaffield, 2011). La evaluación integrada y la evaluación compartida también responden a este modelo.

Por último, para que estas prácticas centradas en el proceso sean efectivas es necesario utilizar técnicas de naturaleza descriptiva y cualitativa, como las rúbricas de evaluación en las que se detallan los indicadores que conducen a una resolución de calidad, para así asegurar que el alumno se regule con autonomía (Díaz, 2006). Por ello, a raíz también de este cambio, encontramos términos como evaluación del desempeño –*performance assessment*– o evaluación basada en evidencias –*evidence-based assessment*–, que se enfocan en señalar la importancia de aquello que es representativo del aprendizaje, es decir, la ejecución. Son aquellas que utilizan trabajos de investigación sobre un tema o para resolver una pregunta que,

de forma recurrente, culminan con la elaboración de un producto final. La evaluación auténtica - *authentic assessment*- es otro modelo que responde a este conjunto de premisas de la evaluación formativa, de la evaluación del, y para el aprendizaje y a través del análisis del desempeño de los alumnos, añadiendo además la dimensión de autenticidad. Esta modalidad es central en nuestra investigación y la desarrollaremos ampliamente en el siguiente apartado.

Como hemos señalado, los modelos de evaluación se conectan e interrelacionan entre ellos, de forma que ninguno en su individualidad puede, por sí solo, definir la evaluación docente de forma exclusiva. Por ejemplo, no podemos decir, para explicar cómo es nuestra evaluación, que es únicamente sumativa, dado que esta es solo una de sus características. Podría ser una evaluación sumativa pero procesual y basada en actividades de autoevaluación. Lo cual reconfigura nuestra concepción inicial de lo que significa sumativa, altamente asociada a un tipo de evaluación final y realizada exclusivamente por el profesor. De hecho, muchos de los docentes sostienen un planteamiento híbrido de prácticas que responden a diversas funciones, de forma que es muy difícil encontrar un sistema de evaluación completamente coherente.

En síntesis, desde nuestra visión, no es tan importante definir el tipo de evaluación sino cómo se caracteriza en su totalidad dado que las opciones son múltiples y diversas. Al respecto, sí encontramos varios patrones que son comunes en la caracterización de la evaluación o, al menos, una evolución clara que va de una evaluación simple, numérica y acreditativa a una compleja, cualitativa y orientada a mejorar la enseñanza y el aprendizaje. Concluimos, pues, que las características de la evaluación han ido evolucionando de un enfoque denominado tradicional a otro enfoque llamado alternativo. Es importante, entonces, profundizar en cuáles son los paradigmas evaluativos y su transformación.

Rueda y Torquemada (2008), en el estudio en el que analizan las concepciones que docentes y estudiantes tienen sobre evaluación, definen en primer lugar cuáles son los paradigmas sobre evaluación predominantes y en constante evolución. Dentro de estos, podemos diferenciar dos conjuntos, un primer grupo de

paradigmas relacionados con la evaluación tradicional, y un segundo grupo de paradigmas conectados con la evaluación alternativa.

En el primer grupo, encontramos el paradigma de la "intuición pragmática", orientado a certificar el éxito de los aprendizajes desde una función social y acreditativa, siendo una evaluación en absoluto formativa; el paradigma de la "docimología o ciencia de los exámenes", centrada en analizar la fiabilidad de los instrumentos y justificar los resultados a raíz del diseño de los mismos, especialmente a través pruebas que poco dan cuenta de las competencias adquiridas; el paradigma de la "evaluación al servicio de la decisión", en la que se implican diferentes voces para tomar la mejor decisión, pero que se aleja del proceso de aprendizaje centrándose en el objetivo.

En el segundo grupo, el paradigma basado en el modelo "tayloriano", centrado en la comprobación de que los resultados alcanzados responden a los objetivos preestablecidos, con la dificultad que comporta describir en términos medibles los comportamientos esperados; el paradigma "la evaluación formativa en la enseñanza diferenciada", dirigido a analizar las necesidades de aprendizaje surgidas durante el proceso, que conlleva una dedicación muy grande en definir el conjunto de informaciones que se van a recoger para juzgar las estrategias a enseñar que orienten a la consecución de los mejores resultados; y el paradigma de "evaluación comprensiva", en el que las evaluaciones deben ayudar a analizar y mejorar sus procesos desde una visión sistémica de lo que ocurre, con la dificultad de invertir una cantidad de tiempo notable en recoger toda la información.

Todos ellos tienen unas finalidades soterradas que en muchos casos comparten varios paradigmas, como también unas deficiencias que otros paradigmas han intentado suplir. Con todo, cada uno a su manera, ha nutrido a través de su discurso las prácticas de evaluación y la consideración de su finalidad dentro de la escuela. Obviamente los últimos son los que han aportado más literatura y experiencias en el campo educativo de la evaluación de los aprendizajes y en muchas de las prácticas actuales podemos observar la manera en que estos paradigmas se conjugan en instrumentos, decisiones y valoraciones variadas. Situémonos,

entonces, en una evaluación formativa y comprensiva, bajo el paraguas de la evaluación alternativa

1.1.1.2. Desde la evaluación alternativa.

De entre todas las clasificaciones expuestas, aquella que históricamente ha sido más nombrada en la literatura sobre evaluación es la conceptualización acuñada por Scriven (1967) en la que se diferencia entre una evaluación sumativa y una evaluación formativa. El nuevo término representó un gran cambio puesto que hace referencia a una nueva función de la evaluación, en la que más allá de servir para medir cuantitativamente los resultados, tal y como se había hecho hasta el momento, la evaluación debe ser un medio para recoger información cualitativa que sirva para mejorar los resultados. Se inició así un nuevo camino que ofreció una alternativa a la evaluación tradicional y que, además, condujo a la proliferación de nuevos términos sobre evaluación, como hemos señalado anteriormente.

La evolución de los paradigmas de evaluación y de los términos sobre evaluación, nos muestra también esta dirección: la eclosión de un conjunto de modelos centrados en el aprendizaje y al servicio de este. El nuevo cambio se define por la necesidad de transitar desde un enfoque de evaluación tradicional a un enfoque de evaluación alternativa (Mateo y Martínez, 2008). Según Álvarez (2001), la evaluación tradicional está definida desde la racionalidad técnica, enfocada a una acción estratégica y dirigida a mantener el *statu quo*. Por contra, la evaluación alternativa está definida desde la racionalidad práctica, enfocada a una acción comunicativa y dirigida a actuaciones críticas y reflexivas. Al primer enfoque van asociados instrumentos de medición de resultados como son los exámenes usados únicamente por el profesor al final de una unidad de enseñanza y, al segundo, cualquier instrumento que permita valorar diferentes evidencias en momentos variados y con la participación del alumnado, entre otros muchos cambios. Mateo (2000) resume muy bien las características diferenciadoras de los dos enfoques distinguiendo entre los instrumentos utilizados, el resultado de la evaluación, su generalización, los datos que se recogen, los referentes de valoración y la participación del alumno (ver Tabla 2).

Tabla 2*Características de la evaluación tradicional y de la evaluación alternativa*

Evaluación tradicional	Evaluación alternativa
Se utilizan mayoritariamente pruebas de elección múltiple, de emparejamiento, de verdadero o falso, de ordenamiento, de completar, etc.	Se realizan experimentos, debates, portafolios, productos de los estudiantes, etc.
Permite al evaluador presentar el conocimiento del alumno mediante una puntuación.	Habilita al evaluador a crear una historia evaluativa respecto del individuo o del grupo.
Tiende a ser generalizable.	Tiende a ser idiosincrásica.
El juicio evaluativo se basa en la recogida de información objetiva y en la interpretación de las puntuaciones.	El juicio evaluativo se basa en la observación, en la subjetividad y en el juicio profesional.
Se focaliza fundamentalmente sobre la puntuación del alumno en comparación con la puntuación de los otros alumnos.	Se focaliza de manera individualizada sobre el alumno a la luz de sus propios aprendizajes.
Provee la información evaluativa de tal forma que inhibe la acción curricular o instruccional.	Provee información evaluativa de manera que facilita la acción curricular.
Coloca la evaluación bajo la decisión del profesor u otra fuerza externa.	Permite a los estudiantes participar en su propia evaluación.

Nota. Extraído de Mateo (2000, p. 67)

Como se puede observar en las características señaladas, estos dos enfoques representan una dualidad que en la práctica determina la manera general de situarse ante la evaluación e implementarla. Por tanto, es necesario conocer este paraguas teórico para poder transformar la evaluación. Además, reflexionar sobre evaluación alimenta y provoca el cuestionamiento respecto a la enseñanza y el aprendizaje, y a la inversa. De hecho, la transición de un modelo evaluativo tradicional a uno alternativo responde también a la evolución de las teorías del aprendizaje y la enseñanza, de forma que las nuevas concepciones pedagógicas son un detonante para el cambio en la evaluación.

Así pues, la relación entre cómo se entiende la enseñanza y el aprendizaje y cómo se conceptualiza la evaluación es ineludible (Lukas y Santiago, 2009) y, por ello, para cambiar la evaluación, también debemos conocer cómo nos posicionamos educativamente. En un primer lugar, es necesario considerar las nuevas teorías del aprendizaje basadas, principalmente, en el enfoque socioconstructivista del aprendizaje. En la Tabla 3, exponemos las principales implicaciones que las teorías

del aprendizaje tienen para la evaluación, según Cabrera (2000, citado en Lukas y Santiago, 2009).

Tabla 3

Cambios que debe asumir la evaluación para responder a las nuevas teorías sobre aprendizaje

Teorías del aprendizaje	Implicaciones para la evaluación
El conocimiento es algo que se construye. El aprendizaje es un proceso de creación de significados a partir de los conocimientos previos: es un proceso de transformación de las estructuras cognitivas del alumnado como consecuencia de nuevos conocimientos.	<ul style="list-style-type: none"> - Promover acciones evaluativas que pongan en juego la significatividad (funcionalidad) de los nuevos aprendizajes. - Evitar los modelos memorísticos en los que sólo se pone de manifiesto la capacidad para reconocer o evocar. - Promover actividades y tareas de evaluación que tengan sentido para el alumnado.
Hay variedad en los estilos de aprendizaje, la capacidad de atención, la memoria, el ritmo de desarrollo y las formas de inteligencia.	<ul style="list-style-type: none"> - Promover distintas formas de evaluación. - Dar oportunidades para revisar y repensar. - Utilizar procedimientos que permitan al alumnado aprender a construir su forma personal de aprender, a manejar autónomamente procedimientos de evaluación y corregir los errores que pueda detectar.
Las personas tienen una ejecución mejor cuando conocen la meta, observan modelos y saben los criterios y estándares que se tendrán en cuenta.	<ul style="list-style-type: none"> - Promover que el alumnado haga suyos los objetivos del aprendizaje y los criterios que se van a utilizar para evaluarlos. - Proporcionar una amplia gama de modelos de ejemplo sobre trabajos de los alumnos para discutir sus características. - Discutir los criterios que se utilizarán para juzgar la ejecución.
Se reconoce que el conocimiento y la regulación de los propios procesos cognitivos son la clave para favorecer la capacidad de aprender a aprender. Es importante saber manejar su propio proceso de aprendizaje.	<ul style="list-style-type: none"> - Promover la autoevaluación. - Estimular los procesos de coevaluación entre el profesor y el alumno y entre éstos entre sí. - Ofrecer retroalimentación tanto de los errores como de los aciertos. - Hacer tomar conciencia al alumnado de lo que han aprendido, así como de los procesos que le han permitido construir nuevos aprendizajes.
La motivación, el esfuerzo y la autoestima afectan al aprendizaje y la ejecución de la persona.	<ul style="list-style-type: none"> - Atribuir los fracasos a razones temporales y externas y los éxitos a razones internas y perdurables. - Estimular relaciones entre el esfuerzo y los resultados. - Incorporar de manera natural tareas de evaluación durante el proceso de enseñanza aprendizaje que puedan servir al alumnado para tomar conciencia de lo que ha aprendido y de las dificultades o lagunas que todavía tiene.

Nota. Extraído de Cabrera (2000, citado en Lukas y Santiago, 2009, p. 152)

Como observamos, las teorías del aprendizaje actuales tienen implicaciones en la evaluación a nivel de instrumentos, finalidad y agente que evalúa, entre otros. Además, van más allá y reclaman modificar aspectos evaluativos en consonancia con los principios del aprendizaje, como el hecho de compartir y consensuar objetivos y criterios de aprendizaje para que los alumnos puedan autorregularse o promover una atribución de los resultados positiva para favorecer la motivación y la autoestima del aprendiz. Asimismo, uno de los siete principios que deberían orientar la creación de entornos de aprendizaje del siglo XXI, enunciado por la OECD (2010) basándose en la naturaleza del aprendizaje, es que se debe implementar una evaluación formativa que permita al alumnado regular su aprendizaje.

Esto conecta con los enfoques pedagógicos actuales en el sistema educativo que están centrados, especialmente en los países de la Unión Europea, en promover situaciones de aprendizaje que permitan la adquisición de competencias. Este cambio en la enseñanza parte de la idea de que el conocimiento que se aprende no se transfiere de forma directa en la realidad de los alumnos, de forma que es necesario hacer explícita esa transferencia a través de metodologías activas que incluyan la aplicación de los saberes (Liesa, Castelló y Becerril, 2018), lo cual implica una evaluación del desempeño. Además, para el desarrollo de un aprendizaje competencial es necesario fomentar los procesos de autorregulación y co-regulación del alumnado (Monereo, Pozo y Castelló, 2001), lo que a su vez implica una evaluación orientada al aprendizaje.

Las teorías del aprendizaje y el enfoque competencial también se conectan con la construcción de una escuela inclusiva en tanto que también requiere de una actualización del rol docente y de las metodologías de aula. Para conseguir una enseñanza inclusiva en los procesos de enseñanza y aprendizaje, consideramos la diversidad como la pluralidad de perfiles de aprendizaje y la obligación de asegurar la presencia, la participación y el progreso de todos los alumnos (Ainscow, 2012), lo cual incluye las prácticas de evaluación. Para ello, además, se requieren muchas acciones y cambios, entre los que cabe destacar las competencias de los docentes a través de mejorar su formación inicial y continua, conseguir que transformen su

discurso sobre inclusión y que desarrollen nuevos roles, que introduzcan nuevas prácticas y que trabajen colaborativamente en una red de apoyo y de recursos (Duran y Giné, 2017; Jiménez, H. Rodríguez, Sánchez y J. Rodríguez, 2018; Liesa, Castelló, Carretero, Cano y Mayoral, 2012).

En conclusión, la relación entre las actuales teorías y concepciones del aprendizaje y de la enseñanza, que reclaman un enfoque competencial, un aprendizaje significativo y una educación inclusiva, tienen implicaciones en la evaluación, de forma que después de analizarlas podemos afirmar con seguridad que debemos situarnos en un enfoque de evaluación alternativa de y en los procesos de enseñanza y aprendizaje con la participación del alumnado.

Por último, si de verdad queremos un sistema educativo coherente con las nuevas premisas sobre la enseñanza, en el que incluimos la evaluación como práctica de aprendizaje, es necesario diseñar un proceso riguroso en el que se compruebe que realmente lo que se evalúa y cómo se evalúa, representa de forma adecuada aquello que se enseña y cómo se enseña (Jornet, González, Suárez y Perales, 2011). Para ello, debemos conocer cuáles son todas las posibilidades que nos ofrecen los diferentes componentes de la evaluación y cómo organizar un proceso evaluativo coherente con nuestro propósito. En los siguientes puntos analizaremos estos aspectos, orientándolos a una evaluación en y de los procesos de enseñanza y aprendizaje.

1.1.2. Los componentes de la evaluación.

Como hemos visto, han ido emergiendo diferentes conceptualizaciones de la evaluación a la par que las prácticas de evaluación se ampliaban, dando respuesta a más posibilidades según diferentes componentes. En realidad, la evaluación es un conjunto de prácticas que se dirigen hacia un enfoque u otro, pero que deben dar respuesta todas a un conjunto de elementos que la definen. En este punto vamos a exponer cuáles son las decisiones que podemos tomar para construir nuestra evaluación orientada al servicio del aprendizaje para, en el siguiente punto, saber cómo podemos diseñar un proceso de evaluación coherente con este enfoque.

Así pues, debemos conocer cuáles son los elementos que componen la evaluación para caracterizarla adecuadamente (ver Tabla 4). Para ello, debemos responder a para qué evaluamos, qué evaluamos, en base a qué referentes evaluamos, con qué evaluamos, cuándo evaluamos y quiénes evaluamos. Las respuestas a estas preguntas no son únicas y definirán nuestro enfoque de evaluación. Además, están interrelacionadas entre sí, de forma que tomar decisiones respecto a uno de estos elementos definirá los siguientes.

Tabla 4

Componentes que definen la evaluación y preguntas conductoras para definirlos

Componentes	Preguntas conductoras
Las finalidades de la evaluación	¿Para qué evaluamos? ¿Qué propósito tiene la evaluación? ¿Cuáles son las funciones de la evaluación? ¿La evaluación sirve para acreditar los aprendizajes? ¿La evaluación sirve para regular los procesos de enseñanza y aprendizaje?
El objeto de la evaluación	¿Qué evaluamos? ¿Cuál es el objeto de la evaluación en y de los procesos de enseñanza? ¿Qué información/datos debemos recoger? ¿Qué entendemos por resultados? ¿Se debe evaluar el proceso y/o el progreso?
Los referentes de valoración	¿Cómo valoramos? ¿De qué manera podemos analizar la información recogida? ¿En base a qué parámetros lo realizamos? ¿Quién establece estos parámetros? ¿Qué se hace con ellos?
Los instrumentos de evaluación	¿Cómo recogemos la información? ¿Con qué instrumentos lo hacemos? ¿Cómo son estos instrumentos? ¿Qué nos permiten? ¿Cuáles son las técnicas y/o actividades que nos facilitarán su aplicación?
Los momentos de la evaluación	¿Cuándo evaluamos? ¿Antes, durante o después del proceso de enseñanza y aprendizaje? ¿Con qué frecuencia lo hacemos? ¿Evaluamos al instante?
Los agentes de la evaluación	¿Quién evalúa? ¿Es el profesor el único responsable de la evaluación? ¿Los alumnos pueden autoevaluarse? ¿Pueden evaluar los alumnos a sus iguales? ¿Es posible que participen otros agentes en la evaluación?

A continuación, vamos a explicar brevemente cada uno de los elementos, justificando cómo orientar la toma de decisiones hacia una evaluación del y para el aprendizaje.

1.1.2.1. Para qué evaluamos.

¿Para qué evaluamos? ¿Qué propósito tiene la evaluación? ¿Cuáles son las funciones de la evaluación? ¿La evaluación sirve para acreditar los aprendizajes? ¿La evaluación sirve para regular los procesos de enseñanza y aprendizaje? Todas estas son las preguntas iniciales que comparten los docentes cuando empiezan a reflexionar sobre la evaluación y sus finalidades. No es gratuito, un buen conocimiento de las diferentes funciones posibilitará el diseño de un sistema de evaluación que responda a la mejora de la educación. Es decir, que la recogida de información, su análisis y el juicio emitido conduzcan a tomar decisiones que permitan mejorar los procesos de enseñanza y aprendizaje. Como apuntan Coll y Martín (1996):

En ocasiones, el proceso de toma de decisiones versará más directamente sobre la propia acción educativa (entonces sirve como instrumento de regulación de la enseñanza: evaluación de la enseñanza); en otras, versará más directamente sobre el proceso de aprendizaje de los alumnos (entonces sirve como instrumento de regulación del aprendizaje: evaluación del aprendizaje); lo ideal, sin embargo, es que verse simultáneamente sobre el proceso de enseñanza y aprendizaje como un todo, cumpliendo una función de regulación y control de dicho proceso. (párr. 16)

Así pues, la evaluación es una actividad fundamental que aporta información para juzgar la enseñanza y el aprendizaje. Sin embargo, aunque este es el enfoque que queremos que predomine, relacionado con la denominada anteriormente evaluación formativa, también cumple otras funciones que suelen ser más recurrentes en las prácticas reales de las escuelas, relacionadas con la evaluación sumativa. Estamos hablando del uso que se le da a la evaluación para tomar decisiones de carácter social, mediante las cuales se demuestran y comunican los resultados obtenidos. Veamos con mayor profundidad cuáles son exactamente las funciones que la literatura sobre la evaluación nos señala como las más importantes (Brown y Remesal, 2012; Coll y Remesal, 2009; Remesal, 2011; Sanmartí, 2010b):

- Una función formativa: de carácter pedagógico y orientada a regular los procesos de enseñanza y aprendizaje. Consiste en que los docentes analizan la enseñanza realizada y sus resultados para tomar decisiones que les permitan acompañar mejor a los alumnos.
- Una función formadora: de carácter pedagógico y orientada a regular los procesos de enseñanza y aprendizaje. Consiste en la que los alumnos analizan su aprendizaje y sus resultados para tomar decisiones que les permitan mejorar, con la guía de los docentes.
- Una función de rendición de cuentas: de carácter social y orientada a la certificación de los aprendizajes. Consiste en dar cuenta a otros agentes educativos y a la sociedad de los resultados obtenidos, con el fin de demostrar la eficiencia del sistema educativo.
- Una función acreditativa: de carácter social y orientada a la certificación de los aprendizajes. Consiste en cuantificar los aprendizajes para diferenciar grados que faciliten orientar o clasificar, para permitir el acceso a otros niveles de enseñanza y para justificar la adquisición de la capacitación necesaria para desarrollar otras actividades.

Consideramos que estas funciones de la evaluación son todas, en mayor o menor medida, necesarias, en tanto que responden a propósitos diferentes pero complementarios. Ninguna de ellas puede obviarse porque las cuatro se encuentran al servicio de la evaluación educativa. Tan necesario es evaluar los aprendizajes para enseñar mejor, por ejemplo, como para decidir si los alumnos están o no capacitados para ciertas actividades. Ahora bien, una evaluación centrada exclusivamente en su carácter más social, donde únicamente se miden y juzgan los aprendizajes, no proporciona suficiente información, ni a docentes ni a alumnos, sobre cómo mejorar. Y la realidad es que tradicionalmente este enfoque y los instrumentos y prácticas que se le asocian ha tenido mucho más peso en la evaluación. De hecho, todavía hoy en día el boletín de notas sigue siendo el protagonista y la mayoría de las actividades de evaluación suelen centrarse en cuantificar los aciertos para poner una nota final y, con ello, juzgar las capacidades de los alumnos determinando con ello, en muchos casos, hasta su futuro profesional.

Si queremos una evaluación al servicio de la educación, debemos dirigir nuestros esfuerzos hacia sus funciones de carácter pedagógico, centrándonos en conocer qué sucede durante el aprendizaje para tomar decisiones educativas y no solo acreditativas. De otra manera, podríamos afirmar que no estamos realizando una evaluación realmente educativa, en tanto que los juicios que emitimos no nos permiten valorar cómo de eficaz está siendo la enseñanza para los alumnos. Hacerlos partícipes para que aprendan a regular su propio aprendizaje. Sin embargo, este nuevo enfoque representa un reto para la gran parte del colectivo de profesores, los cuales han experimentado, mayoritariamente, durante su etapa escolar y formativa una evaluación acreditativa.

Al respecto, es interesante indagar cuáles son las creencias sobre evaluación que sostiene el profesorado. En su estudio, Brown (2006, 2006) diseñó y validó un instrumento para evaluar las concepciones sobre evaluación que tenían los docentes en relación con las finalidades de la evaluación (*Teachers' Conceptions of Assessment-III*), distinguiendo las siguientes dimensiones: la evaluación dirigida a rendir cuentas al sistema, la evaluación enfocada a acreditar al alumnado, la evaluación al servicio de la mejora de la educación y la evaluación como un aspecto irrelevante. Partiendo de este marco se realizaron otros estudios para profundizar en diferentes aspectos como el nivel de acuerdo entre profesores en formación y profesores en ejercicio (Brown y Remesal, 2012) o la definición de las concepciones de profesores en contextos de educación obligatoria marcados por diferentes políticas educativas (Brown, Hui, Yu y Kennedy, 2011; Brown y Michaelides, 2011). En sus resultados destaca una predominancia del enfoque de evaluación sumativa, es decir, una evaluación cuya finalidad se orienta a acreditar el aprendizaje y justificar las prácticas educativas.

Otro punto interesante sobre este tema es analizar por qué o cómo se construyen en los profesores determinadas creencias en base a la diversidad de funciones. Coll y Remesal (2009) realizaron un estudio en base a las finalidades interpretándolas no como excluyentes y dicotómicas, sino posicionándolas dentro de un continuo que va de un polo social o a uno pedagógico, considerando la complejidad que supone cumplir con todas. En los resultados señalan que los profesores de educación secundaria tendían más a un polo social, mientras que

entre los maestros de educación primaria primaban las funciones pedagógicas. Este resultado lo relacionan con la estructura propia de la etapa secundaria, más centrada en acreditar para poder acceder a niveles educativos superiores. Un estudio posterior (Remesal, 2011) señala que las concepciones de los docentes sobre las funciones de la evaluación podrían estar condicionadas por la influencia que ejercen las demandas de evaluación externas.

Es importante, continuar ampliando el estudio sobre las concepciones que tienen los docentes sobre las funciones de la evaluación y cómo estas han evolucionado, dado que pueden llegar a determinar la práctica evaluativa de los docentes y el conocimiento nos puede ayudara a orientar las formaciones sobre evaluación (Barnes, Fives y Dacey, 2017; Levy-Vered y Nasser-Abu, 2018; Lumsden y Michaelides, 2011; Muñoz, Palacio y Escobar, 2011; Segers y Tillema, 2011).

En conclusión, las funciones de la evaluación son todas legítimas, pero el sistema, que tradicionalmente ha reforzado sus propósitos acreditativos y de rendición de cuentas, relacionados con la evaluación sumativa, debe centrarse en promover sus propósitos de regulación del aprendizaje, relacionados con la evaluación formativa. Los profesores, que ahora cuentan con el apoyo de un sistema educativo más centrado en el aprendiz y en el desarrollo de sus competencias, deben transitar de unas creencias a otras con la ayuda de nuevas prácticas de evaluación. Para ello, deben determinar cuidadosamente qué, cómo, quién y cuándo se debe evaluar, para conseguir dirigirse hacia una evaluación de y en los procesos de enseñanza y aprendizaje. En los siguientes puntos vamos a analizar estos aspectos que conforman la evaluación.

1.1.2.2. Qué evaluamos.

¿Qué evaluamos? ¿Cuál es el objeto de la evaluación en y de los procesos de enseñanza? ¿Qué información/datos debemos recoger? ¿Qué entendemos por resultados? ¿Se debe evaluar el proceso y/o el progreso? Definir el qué de la evaluación, cuál es la información que vamos a recoger para tomar decisiones, no es un proceso fácil, sin embargo, sí es un aspecto altamente determinante para concretar nuestra evaluación. Los datos que queramos recoger condicionarán los

instrumentos de evaluación, el momento y los agentes que participarán, entre otros, de tal forma que si no hay una coherencia entre ellos podremos devaluar el poder formativo de algunas estrategias muy eficaces. Por ejemplo, para corregir una redacción y analizar la cantidad de faltas de ortografía, no tiene mucho sentido utilizar una rúbrica para indicar si hacer dos, tres o diez faltas está mejor o peor. Así pues, para elegir cómo vamos a evaluar primero debemos saber qué estamos evaluando.

En el caso de los procesos de enseñanza y aprendizaje, podemos decir que nuestro objeto de evaluación son sus aprendizajes y, de forma extensible las actuaciones del profesorado que lo hacen posible. Ahora bien, la palabra aprendizaje es bastante amplia y debemos operativizar para concretar. Por un lado, podemos evaluar contenidos y, por otro lado, los procesos de adquisición de dichos contenidos. Cuando decimos contenidos, nos referimos a aquello a lo que se dirige el acto de enseñanza y aprendizaje, sean conocimientos, habilidades o competencias. Cuando decimos procesos de adquisición, nos referimos al momento en qué se encuentra el alumno con relación a ese aprendizaje y cuál es su progreso, la distancia que hay entre su punto de partida y conocimientos previos, hasta su momento y conocimientos actuales. Vamos, pues, a clarificar estos términos.

Nos vamos a enfocar primero en la definición sobre qué entendemos por contenidos y por competencias. En el caso español, según Ley Orgánica 8/2013, de 9 de diciembre, para la mejora de la calidad educativa (LOMCE), en su Real Decreto 126/2014, de 28 de febrero, por el que se establece el currículo básico de la Educación Primaria, se definen los contenidos y las competencias en el artículo 2 de la siguiente manera:

Los contenidos o conjuntos de conocimientos, habilidades, destrezas y actitudes que contribuyen al logro de los objetivos de cada enseñanza y etapa educativa y a la adquisición de competencias. Los contenidos se ordenan en asignaturas, que se clasifican en materias, ámbitos, áreas y módulos en función de las enseñanzas, las etapas educativas o los programas en que participe el alumnado.

Competencias: capacidades para aplicar de forma integrada los contenidos propios de cada enseñanza y etapa educativa, con el fin de lograr la realización adecuada de actividades y la resolución eficaz de problemas complejos. (Artículo 2)

En relación con los contenidos, la evaluación tradicional se ha centrado en la valoración de los conocimientos, a través de actividades en las que los alumnos debían reproducir los saberes transmitidos en clase. La propia agrupación de los contenidos, en temas y en asignaturas, unido a la visión sumativa de la evaluación y a la importancia que socialmente se le ha dado con anterioridad a memorizar conocimientos, ha ayudado a perpetuar estas prácticas de evaluación. Progresivamente, esta visión se fue ampliando con la idea de que también era importante evaluar las habilidades y estrategias de los alumnos, es decir, su proceso de aprendizaje.

Así pues, también debemos valorar las estrategias de aprendizaje del alumno, definidas como la toma de consciencia del proceso seguido, la habilidad para planificarlo, regularlo y evaluarlo. Según Monereo (1997) se trate de enseñar a los alumnos “a reflexionar sobre su propia manera de aprender, ayudándoles a analizar las operaciones y decisiones mentales que realizan, con el fin de mejorar los procesos cognitivos que ponen en acción” (p. 8). Y, para ello, debemos realizar actividades de evaluación en las que tengan que explicar y explicitar cuál ha sido su proceso. Esta nueva perspectiva marcó el camino para una evaluación de competencias.

En relación con las competencias, podemos afirmar que este es un término mucho más complejo y difícil de operativizar para evaluarlo. Las competencias se caracterizan por estar constituidas por la activación de saberes conceptuales, procedimentales y actitudinales, por tomar sentido durante la acción cuando esta movilización de saberes se pone al servicio de la resolución de un problema y por poderse desarrollar de forma progresiva en diferentes contextos y experiencias vitales (Ion y Cano, 2012). Por tanto, para evaluar competencias requeriremos escenarios especiales y una situación realista donde la competencia se pueda desarrollar (Díaz, 2006).

Además, según De la Orden (2011) existen dos interpretaciones del concepto de competencia: una perspectiva que se centra en evaluar la integración y movilización de los saberes para resolver problemas prototípicos y otra perspectiva que se centra más en la participación real en un contexto determinado de forma eficaz. Por tanto, desde estas perspectivas no es suficiente con plantear actividades de evaluación para saber si resuelven un problema con los conocimientos y habilidades que tienen, sino que, además, debemos analizar cómo lo resuelve para saber en qué nivel a desarrollado la competencia. Para ello se necesita una creación de escenarios para la puesta en acción (Tonda y Medina, 2013). Esta nueva perspectiva ha reforzado que la evaluación tenga como objeto también el proceso y, especialmente, el progreso de los aprendices.

Este último aspecto, es seguramente uno de los más difíciles de evaluar. Consiste en reconocer el aprendizaje del alumno con relación a sí mismo, a sus avances y a la consecución de sus objetivos personales; está relacionado con una evaluación criterial y con una evaluación inclusiva. En ocasiones, encontramos alumnos que no alcanzan los estándares de aprendizajes esperados, pero que, en relación con sus conocimientos y habilidades previas, han mostrado una mejora significativa y en algunos casos han hasta aprendido mucho más que los que sí han alcanzado los estándares. Este dilema es común a muchos docentes y no hace más que resaltar la importancia de la evaluación como actividad de aprendizaje. Una actividad que se debería centrar en ayudar a reconocer los avances y no tanto en certificar los resultados.

En conclusión, si nos situamos en un enfoque de enseñanza competencial y evaluación orientada al aprendizaje, deberemos evaluar el desarrollo de las competencias a través del análisis de procesos de resolución y el progreso del alumno, sin dejar de banda la valoración sobre la adquisición de los conocimientos, habilidades y estrategias necesarias para actuar de forma competente. Pero, es necesario ir más allá y determinar en relación con qué vamos a evaluar estos contenidos y competencias y estos procesos y progresos del alumnado. En el siguiente punto vamos a definir esta cuestión: los referentes para juzgar y tomar decisiones.

1.1.2.3. En referencia a qué evaluamos.

¿Cómo valoramos? ¿De qué manera podemos analizar la información recogida? ¿En base a qué parámetros lo realizamos? ¿Quién establece estos parámetros? ¿Qué se hace con ellos? La respuesta la encontramos en tomar decisiones previas sobre cómo medir y valorar la información que recogemos. Es decir, qué aspectos darán cuenta de que tanto los contenidos como las competencias han sido aprendidos, desarrollados y en qué grado. Para ello es necesario determinar una referencia consensuada con la que poder comparar la ejecución, la demostración del aprendizaje.

Este aspecto es comúnmente olvidado cuando se piensa en evaluar, sin embargo, sin el establecimiento de unos referentes previos para la evaluación ésta no se produce. Según Lukas y Santiago (2004, p. 100), "para que se dé la evaluación se deben emitir juicios de valor sobre el objeto evaluado. Para ello, la información recogida sobre los criterios preestablecidos será valorada teniendo en cuenta los referentes señalados como indicadores de éxito." Al respecto, tradicionalmente, ha habido dos clases de valoración: la normativa y la criterial. En la evaluación de referencia normativa se comparan los resultados de un alumno con los resultados del resto de alumnos de un grupo determinado. Los instrumentos surgidos de esta clase de valoración miden lo que los estudiantes conocen en un momento determinado, pero no sus aprendizajes en términos absolutos. En la evaluación de referencia criterial se establece una referencia previa respecto a la cual se juzga el nivel de logro del alumno (Mateo, 2000).

Nos posicionamos en la valoración basada en criterios como promotora de una evaluación más formativa. Dentro de ésta distinguimos diversos conceptos como los estándares de aprendizaje evaluables, los criterios de evaluación, los indicadores de evaluación, los criterios de calificación y los requisitos de la evaluación. Conviene diferenciarlos para por situarnos en una evaluación realmente formativa y compartida y no utilizar estrategias de forma contradictoria (López-Pastor y Pérez-Pueyo, 2017).

Empezando por los estándares y los criterios, actualmente, en el contexto español, según Ley Orgánica 8/2013, de 9 de diciembre, para la mejora de la calidad educativa (LOMCE), en su Real Decreto 126/2014, de 28 de febrero, por el que se establece el currículo básico de la Educación Primaria, se definen de la siguiente manera:

Estándares de aprendizaje evaluables: especificaciones de los criterios de evaluación que permiten definir los resultados de aprendizaje, y que concretan lo que el alumno debe saber, comprender y saber hacer en cada asignatura; deben ser observables, medibles y evaluables y permitir graduar el rendimiento o logro alcanzado.

Criterios de evaluación: son el referente específico para evaluar el aprendizaje del alumnado. Describen aquello que se quiere valorar y que el alumnado debe lograr, tanto en conocimientos como en competencias; responden a lo que se pretende conseguir en cada asignatura. (Artículo 2)

Ambos son una especificación de los niveles de desempeño que debe mostrar el alumno que domina una competencia; en el caso de los criterios, son los aspectos específicos para la observación de la calidad del aprendizaje y, en el caso de los estándares, son la operativización de los niveles de calidad del aprendizaje (Jornet et al., 2011). Los estándares ayudan a realizar la evaluación acreditativa y los criterios a realizar una evaluación pedagógica. De forma específica, en la concreción de los criterios podemos encontrar de dos tipos (Sanmartí, 2010b): los criterios de realización que especifican qué aspectos deben evaluarse, y los criterios de resultado, que fijan los diferentes grados en los que se pueden definir los criterios de realización. En algunos estudios, también se denominan indicadores de logro, siendo un enunciado en el que se concreta aquello en lo que se desgrana el criterio, de manera que se ayuda al evaluar a centrar la observación.

Por otro lado, conviene no confundir estándares, criterios e indicadores de evaluación con los criterios o parámetros de calificación o los requisitos de ésta, conceptos también ampliamente utilizados. Los parámetros de calificación se utilizan normalmente para cuantificar el resultado del aprendizaje y se refieren a la porcentualización de las actividades. Y, los requisitos de la evaluación se

establecen cuando definimos cuáles son las condiciones específicas para lograr acceder a los criterios de calificación, por ejemplo, asistencia, entrega de actividades, participación en clase, etc. El uso de estos parámetros y condiciones está asociado a una evaluación sumativa en la que se pretende acotar al máximo para que ninguna variable pueda condicionar la supuesta objetividad de los métodos más tradicionales de medición.

Dentro de un enfoque de evaluación orientada al aprendizaje, establecer criterios de evaluación es de vital importancia para conseguir que las actividades evaluativas proporcionen soporte a los alumnos para aprender a regularse, reflexionar y mejorar sus aprendizajes. Para ello, es necesario que los alumnos participen en la definición de los criterios y en el uso de estos durante el proceso de aprendizaje. Desarrollaremos esta idea en profundidad cuando hablemos del diseño del proceso de evaluación orientada al aprendizaje. Por último, también debemos señalar que la totalidad de los miembros del equipo educativo de los centros deben responsabilizarse, en base al currículum, de seleccionar y distribuir cuándo y de qué manera se deben alcanzar dichos estándares y cómo se van a concretar y especificar los criterios para evaluarlos. De forma que se consiga una coherencia vertical y horizontal en la programación de la evaluación. Además, el discurso sobre estos criterios promoverá el debate sobre de qué manera se debe evaluar, aspecto que desarrollamos a continuación.

1.1.2.4. Con qué evaluamos.

¿Cómo recogemos la información? ¿Con qué instrumentos lo hacemos? ¿Cómo son estos instrumentos? ¿Qué nos permiten? ¿Cuáles son las actividades que nos facilitarán su aplicación? Las respuestas, como en los casos anteriores, no son unívocas ni cerradas. Recordemos que para evaluar se necesita recoger datos y luego analizarlos según los criterios de evaluación establecidos. Por lo tanto, deberemos definir tanto el instrumento y/o actividad que utilicemos para recoger esos datos y el que utilizaremos para valorarlos. Por ejemplo, en educación infantil, realizar una observación detallada sería una actividad de evaluación y, el instrumento, el documento de soporte donde registremos esta observación. Ahora bien, para valorar si esa formación se ajusta o no a los criterios de valoración,

debemos juzgar si los datos que tenemos responden a lo que queríamos y de qué manera, es decir, discriminar la información relevante y analizarlos. Para ello necesitamos también un instrumento de valoración, como una tabla donde organicemos el listado de criterios e indiquemos si se cumplen por alumnos.

Así pues, mientras el instrumento o la actividad permitan recoger información o analizarla, cumplirá su función. Si queremos utilizarlo para recoger información que los alumnos deben transmitirnos, ha de ajustarse a los objetivos programados, a la complejidad cognitiva que se quiere evaluar y a las necesidades de los diferentes alumnos. Si queremos utilizarlo para valorar información, debe contener los criterios y/o indicadores de calidad que utilizaremos; en el caso que sean los alumnos los que los utilicen para evaluarse, deberán ser comprensibles a su edad madurativa y a su conocimiento sobre qué significa evaluarse.

Además, un mismo instrumento o actividad puede utilizarse varias veces durante la programación didáctica o proyecto de investigación de diferentes formas. Al respecto, una enumeración muy clarificadora es la de Sanmartí (2007), quien expone una tabla de posibles instrumentos a utilizar según la función que cumplan y el momento en el que se apliquen, antes como diagnóstico, durante como regulación y final como calificación (ver Tabla 5).

Tabla 5
Inventario de instrumentos de evaluación según los objetivos de evaluación

Objetivo	Instrumentos y actividades
Diagnos is o evaluación inicial	<ul style="list-style-type: none"> - Preguntas abiertas o muy contextualizadas (analizadas intentando comprender la lógica del que aprende, por ejemplo, por medio de redes sistémicas) - Preguntas de opción múltiple diseñadas a partir de las respuestas obtenidas en preguntas abiertas. - Mapas conceptuales. - Formularios KPSI (Knowledge and Prior Study Inventory). Posibilitan identificar lo que el alumno cree que sabe y no tanto lo que realmente sabe, pero a veces la tarea del profesor es ayudarle a reconocer que su autoevaluación inicial no es adecuada - Conversación libre, pero prestando mucha atención a los que no intervienen. - Portafolio o carpetas de trabajo

<p>Evaluación de la representación de los objetivos</p>	<ul style="list-style-type: none"> - Explicitación de objetivos por parte del docente y autoevaluación de los alumnos acerca de cuándo creen que se van apropiando de ellos. - Verbalización de objetivos por parte de los alumnos expresando qué creen que están aprendiendo o han aprendido - Diarios de clase que recogen las opiniones de los alumnos sobre qué creían que iban a aprender en contraste con lo que creen que han aprendido. - Mapas conceptuales que se van construyendo colectivamente (o analizando si ya se ha construido) a medida que se aprende un nuevo concepto y se relaciona con los demás. - Formularios KPSI (si se pregunta sobre el grado de conocimiento de los objetivos de aprendizaje previstos y, posteriormente, van revisando si creen que han mejorado su comprensión). - Análisis de trabajos de alumnos de cursos anteriores y explicitación, por parte de los que aprenden, de las características de la “producción” que creen que se espera que sean capaces de realizar al final del proceso de enseñanza.
<p>Evaluación de la calidad de anticipación y planificación de la acción</p>	<ul style="list-style-type: none"> - Bases de orientación, cartas de estudio, cartas de navegación, en todas ellas los alumnos deben resumir las operaciones que creen necesarias para realizar un determinado tipo de tarea. - Diarios de clase que recojan qué creen que han aprendido y cómo lo han hecho. - Mapas conceptuales que organicen e interrelacionen los conceptos aprendidos. - Esquemas, diagramas de flujo... - V de Gowin, en las que tengan que establecer relaciones entre los referentes teóricos y lo que hacen para ejecutar una tarea - Resúmenes de lo que han aprendido
<p>Evaluación de la representación o percepción de los criterios de evaluación</p>	<ul style="list-style-type: none"> - Elaboración de parrillas de criterios de evaluación por parte de los alumnos, que luego aplicarán a la autoevaluación de sus producciones. - Aplicación de parrillas de criterios de evaluación elaboradas por el profesorado. - Contratos de evaluación, a través de los cuales se negocian y se pactan los contenidos y criterios de evaluación. - Diarios de clase que hagan referencia a qué creen que han aprendido bien y qué no acaban de entender o de saber hacer. - Elaboración, por parte de los alumnos, de preguntas para un posible examen final y concreción de los criterios de evaluación. - Elaboración y corrección de ejercicios en actividades de coevaluación. - Revisión de las respuestas a preguntas o actividades planteadas en la evaluación inicial, explicitando los criterios aplicados para valorarlas.
<p>Evaluación de los resultados de los aprendizajes realizados</p>	<ul style="list-style-type: none"> - Pruebas escritas que incluyen preguntas “productivas”, es decir, aquellas cuya respuesta exija relacionar conocimientos aplicándolos al análisis de situaciones no trabajadas anteriormente. - Mapas conceptuales elaborados por los alumnos. - Realización de proyectos o trabajos que comporten aplicar los conocimientos aprendidos. Pueden concretarse en escritos, maquetas, dramatizaciones, murales... - Exposición oral sobre algún tema en el que haya que demostrar las relaciones con lo aprendido. - Juegos de rol

-
- Responder de nuevo al cuestionario inicial y especificar los cambios producidos en los conocimientos.
-

Nota. Extraído de Sanmartí (2007, pp. 103-107)

Como avanzamos anteriormente, los instrumentos y actividades no son más o menos eficaces por sí solos, sino que depende de nuestros propósitos de evaluación, del objeto de evaluación, de la persona evaluada y del momento en el que se aplique. Lo que sí es cierto es que algunos instrumentos que han tenido más resonancia en la literatura sobre evaluación educativa que otros y algunos otros son más los más utilizados históricamente en las aulas.

Este último caso es el de los tests, exámenes o pruebas de evaluación, los cuales tienen una preponderancia notable en la tradición evaluativa. Este instrumento sirve para recoger información, normalmente de carácter conceptual, y a la vez, para valorar esa información, de forma sencilla y rápida, dado que las opciones de respuestas únicamente son correcto o incorrecto. Hoy en día, las últimas investigaciones sobre, por ejemplo, cómo son los exámenes de educación secundaria en el contexto español, demuestran que mayoritariamente se centran en elementos conceptuales de baja complejidad cognitiva (Gómez y Miralles, 2013; Gómez, Monteagudo y Miralles, 2018; Monteagudo y Vera, 2017;). Afortunadamente, también existen nuevos proyectos que tratan de transformar este instrumento en uno capaz de evaluar competencias (Mejías y Monereo, 2018; Villarroel, Boud, Bloxham, D. Bruna y C. Bruna, 2019), tal y como expondremos en apartados posteriores.

El siguiente grupo de instrumentos, aunque han tenido una menor incidencia, en la actualidad se están revalorizando al resultar idóneos para conseguir una evaluación orientada al aprendizaje y al desarrollo de las competencias. Por un lado, tenemos un conjunto de instrumentos que permiten recoger evidencias y hacer explícito el proceso de aprendizaje y demostrar el desempeño real, como son las carpetas de aprendizaje o portfolios, el diario de aprendizaje, los contratos de aprendizaje, la resolución de problemas complejos, el análisis de casos o la creación de productos específicos.

Estos instrumentos con frecuencia son complementarios entre ellos. Por ejemplo, en las carpetas de aprendizaje se incluyen evidencias de muchos de los instrumentos que hemos enseñado en la tabla como los mapas conceptuales, los cuestionarios, etc. Otro ejemplo, para la creación de productos se suelen utilizar instrumentos intermedios para guiar la resolución ya trabajados con anterioridad como las tablas metacognitivas. Lo importante es que estos instrumentos permiten evaluar competencias complejas (Barberà, 2005; Barberà, Gewerc y Rodríguez, 2016; Díaz y Barroso, 2014; Guash, Guàrdia y Barberà, 2009; Sabirón y Arraiz, 2013).

Por otro lado, encontramos un conjunto de instrumentos que incluyen criterios para valorar diversas evidencias, procesos y ejecuciones, como son las bases de orientación, las tablas de indicadores o las rúbricas de evaluación. Las rúbricas son especialmente significativas porque permiten no solo señalar qué aspectos se valoran sino, además, en qué niveles se pueden lograr y, por tanto, ayudan a comprender cuáles son los siguientes pasos, es decir y autoevaluarse (Ruiz-Teroba, Vázquez y Sevillano-García, 2017). Se pueden hacer rúbricas específicas o de competencias concretas (Alsina, 2013). Además, los alumnos y estudiantes valoran positivamente la explicitación de los criterios de evaluación (Yamila, 2016). Para una evaluación de y en los procesos de enseñanza y aprendizaje, claramente los instrumentos que permiten recoger evidencias y analizarlas de forma autónoma son los más adecuados dentro del paradigma actual.

En conclusión, existen instrumentos que permiten recoger datos y otros analizarlos, algunos con mayor tradición que otra y todos significativos, en tanto que la elección de estos depende del resto de componentes que definan nuestra evaluación. También de las habilidades que cada docente tenga y la predisposición al cambio que demuestre, puesto que la introducción de nuevos instrumentos es compleja y tiene un coste, especialmente si introducimos instrumentos que deben utilizar los alumnos, los cuáles hemos de formar al respecto como veremos en el siguiente punto. Sin embargo, está claro que si queremos una evaluación orientada al aprendizaje es necesario realizar este cambio y la inversión va a merecer la pena.

1.1.2.5. Cuándo evaluamos.

¿Cuándo evaluamos? ¿Antes, durante o después del proceso de enseñanza y aprendizaje? ¿Con qué frecuencia lo hacemos? ¿Evaluamos al instante? Para responder a estas preguntas debemos tener claro que hacemos referencia a la evaluación inserta dentro del proceso de enseñanza y aprendizaje. Obviamente, después de que los alumnos realicen una actividad, podemos corregirla y comunicar los errores, y también podemos evaluar al finalizar un curso escolar. Pero en este punto vamos a situarnos en cómo las actividades de evaluación se distribuyen a lo largo de una programación didáctica o proyecto de investigación, con el fin de dar cuenta sobre cómo se han logrado los objetivos programados y de qué manera pueden mejorar los aprendizajes.

Encontramos pues, tres momentos claves que se corresponden con el antes, durante y después de la programación de enseñanza: la evaluación inicial, la evaluación continua y la evaluación final. Esta diferenciación, emergió de forma paralela al conjunto de reflexiones sobre cómo la evaluación debe tener una función más pedagógica y, por tanto, si se consideran los tres, está asociada a una evaluación formativa. También se relaciona con la evaluación del conocimiento estratégico, en tanto que, por un lado, se centra en evaluar el proceso en diferentes momentos y, por otro, en desarrollar la capacidad metacognitiva del alumnado (Monereo, 2003). A continuación, pasamos a describir los tres momentos de la evaluación señalados.

La evaluación inicial, también denominada evaluación diagnóstica, se realiza al comenzar la programación didáctica o alguna de las actividades de ésta que trabajan contenidos claves, con el objetivo de valorar los conocimientos previos y las habilidades de los alumnos. Esta actividad permite a los alumnos descubrir cuál es el tema y qué saben sobre él, de manera que recuperan sus saberes al respecto para poder utilizarlos. A los profesores les permite ajustar la enseñanza a las necesidades detectadas. Esta evaluación inicial, aportará mayor valor pedagógico si el resto de los componentes, como el objeto o los criterios de evaluación, se enfocan también a una evaluación formativa. Por ejemplo, en una unidad para trabajar los algoritmos matemáticos, si únicamente pasamos una prueba con una

batería de ejercicios de cálculos, descubriremos si saben o no realizar los cálculos o cuántos, pero no cuáles son las dificultades que tienen al respecto ni si son capaces de utilizarlos en la vida real. Por lo tanto, la evaluación inicial puede ayudar a planificar y ajustar la enseñanza y/o a los alumnos su aprendizaje según se conciben el resto de los componentes.

La evaluación continua o continuada, se realiza durante el desarrollo de la unidad didáctica o proyecto investigación. Consiste en realizar diferentes actividades que permitan a los alumnos poner a prueba sus conocimientos y en recoger diferentes evidencias de las actividades que den cuenta de su progreso. Este componente, en la literatura sobre evaluación, ha sido clave para diferenciar entre una evaluación sumativa y una evaluación formativa, a la que claramente propone dar respuesta. Sin embargo, en algunos casos la evaluación continua también se identifica con evaluación sumativa. Por ejemplo, cuando se realizan pruebas parciales sobre lo que saben los alumnos en diferentes momentos, antes del examen final, o cuando se calcula la nota final estableciendo porcentajes obtenidos en diferentes productos de aprendizaje como el dossier, la actitud y los trabajos que se realizan a lo largo del proceso. Pero, aunque se recogen evidencias de forma continuada, al no posibilitar el aprendizaje de los alumnos, no se relaciona con una evaluación formativa.

En otros casos, la evaluación continua sí que responde claramente a una evaluación formativa, en tanto que las prácticas de evaluación que se proponen permiten recoger evidencias que el alumno analizará, se consensuan y utilizan criterios de evaluación para coevaluarse o autoevaluarse y cuando se valora el proceso de resolución, se establecen metas para acompañar el progreso, entre otras acciones. Es decir, cualquier actividad de evaluación que permita a los alumnos tener el control del aprendizaje (Monforte y Farías, 2013) y no únicamente dar cuenta de lo que sabe, y que posibilite al docente experimentar diferentes estrategias y cambiar su planificación inicial para ajustarse a las necesidades de los alumnos. Este enfoque en la evaluación durante el proceso de enseñanza y aprendizaje responde, claramente, a una evaluación orientada al aprendizaje.

En este mismo grupo, también podemos clasificar aquella evaluación que utiliza técnicas para evaluar *in situ*, denominada también evaluación de aula. Es decir, un conjunto de estrategias para valorar la comprensión de los alumnos mientras se enseña o realiza alguna actividad de aprendizaje. El objetivo es ajustar al instante la propuesta de enseñanza utilizando mecanismos para que todos los alumnos participen activamente en la construcción del conocimiento y lo hagan en la dirección adecuada. Algunas de las estrategias son compartir las expectativas de aprendizaje, realizar preguntas específicas que permitan desarrollar la metacognición en la dirección adecuada, aplicar tareas que proporcionen evidencias, dar retroalimentación con comentarios positivos que impulsen a avanzar, ajustar las condiciones como el tiempo o los recursos según los perfiles de los alumnos y permitir y facilitar la relación entre iguales para que compartan sus avances y se ayuden (Sondergeld, Bell y Leusner, 2010). Para que estas técnicas sean efectivas y las utilicen los docentes, deben ser claras de forma que sea posible implementarlas en diferentes cursos y materias, sin un coste muy grande de recursos y tiempo (Leahy, Lyon, Thompon y Wiliam, 2005).

Por último, encontramos la evaluación final, que consiste en aplicar tareas de evaluación una vez finalizada la unidad didáctica o proyecto de investigación. Como en los anteriores momentos, esta evaluación será más o menos sumativa dependiendo del resto de componentes. Sin embargo, en este caso, también se pretende realizar una evaluación que valore todas las evidencias para, además de realizar una comunicación constructiva de los resultados, acreditar los aprendizajes. Atendiendo a las diferentes funciones de la evaluación, se debe emitir una calificación final representativa de la consecución o no de los estándares de aprendizaje esperados y nivelada según los criterios e ítems propuestos, y realizar una retroalimentación que permita a los alumnos y los docentes tomar decisiones sobre qué aspectos se van a mejorar en futuros aprendizajes y futuras programaciones didácticas o proyectos.

En síntesis, los momentos de la evaluación responden a una evaluación estratégica que permite regular los aprendizajes según se pongan al servicio de la toma de decisiones de carácter pedagógico o simplemente se dirijan a certificar los aprendizajes. En este último caso, incluso puede realizarse únicamente una

evaluación final, sin tener en cuenta los conocimientos previos ni el progreso de los alumnos, aunque esta opción no se corresponde para nada con una enseñanza que responda a las perspectivas actuales sobre cómo aprenden los alumnos. Por tanto, podemos asegurar que los tres momentos de la evaluación son claves en el diseño del sistema de evaluación responden a una dinámica de enseñanza estratégica y, sobre todo, determinan que la integración de la evaluación en los procesos de enseñanza y aprendizaje. Partiendo de este objetivo de integración, debemos considerar entonces quiénes son los agentes que participan en la evaluación, tal y como desarrollamos en el siguiente punto.

1.1.2.6. Quiénes evaluamos.

¿Quién evalúa? ¿Es el profesor el único responsable de la evaluación? ¿Los alumnos pueden autoevaluarse? ¿Pueden evaluar los alumnos a sus iguales? ¿Es posible que participen otros agentes en la evaluación? El agente evaluador es la persona que va a participar activamente en las prácticas de evaluación. Para responder a estas preguntas, en primer lugar, debemos diferenciar entre la existencia de lo que se denomina una evaluación interna, en la que son los participantes del objeto de evaluación los que evalúan y una evaluación externa, en la cual unos agentes externos son los encargados de ejecutar la evaluación, especialmente aplicando instrumentos y valorando los resultados (Lukas y Santiago, 2009).

Cuando hablamos de la evaluación interna, identificamos como agentes de la evaluación al alumnado y al profesorado protagonistas de los procesos de enseñanza y aprendizaje. Estos pueden evaluar su propio trabajo y/o el trabajo de los otros, entonces hablamos de autoevaluación o de heteroevaluación respectivamente. Lo más habitual es hablar de heteroevaluación cuando los profesores evalúan a los alumnos, pero ¿quién evalúa al profesor? En este sentido, un aspecto generalmente olvidado es considerar que el profesor puede evaluarse a sí mismo y que los alumnos también pueden, y deben, valorar si les ha funcionado o no la propuesta didáctica realizada por el docente. De hecho, preguntar a los alumnos qué ha funcionado ayuda a los maestros a mejorar sus habilidades de enseñanza (Wiggins, 2011).

Centrándonos en la evaluación de la enseñanza, tanto por parte de los alumnos como de los propios profesores, encontramos que en la literatura sobre evaluación todavía no es este un aspecto muy desarrollado, aunque sí se recoge o menciona en los programas de evaluación de la calidad de la docencia. Existen algunas propuestas al respecto, especialmente en proyectos de investigación sobre la mejora de la docencia universitaria, en los que se hace énfasis en la autoevaluación docente y se utilizan instrumentos como las rúbricas; y proyectos que promueven la reflexión sobre la práctica docente a través del análisis de incidentes críticos como estrategia para mejorar metodologías y competencias docentes (Nail et al., 2012; Monereo, Badia et al., 2009; Monereo, Weise y et al., 2013) y en los que también se consideran las representaciones que los alumnos tienen sobre la enseñanza (Monereo, Panadero y Scartezini, 2013). Es necesario invertir más en este aspecto de forma que la autoevaluación que de manera informal realizan los profesores sobre su práctica, se transforme en una actividad cotidiana y planificada. En este sentido, las propuestas de docencia compartida o co-docencia, pueden favorecer este objetivo (Oller, Navas y Carrera, 2018).

Respecto a la evaluación de los aprendizajes, el docente ha sido siempre el protagonista. Sin embargo, como ya hemos avanzado anteriormente, si queremos una evaluación formativa, los alumnos deben participar en el proceso de evaluación. De esta concepción nacen las actividades en las que el alumno reflexiona sobre su proceso y/o resultado de aprendizaje, denominadas de autoevaluación, y en las que son otros iguales los que participan de esta valoración, denominadas de coevaluación o evaluación entre iguales. Tanto las actividades de autoevaluación como de coevaluación están respaldadas por una extensa literatura al respecto que demuestra los beneficios que ésta tiene para el aprendizaje de los alumnos y los estudiantes, especialmente en tanto que ayuda al desarrollo de la capacidad de autorregulación y la competencia de aprender a aprender (Panadero, Andrade y Brookhart, 2018; Panadero y Alonso-Tapia, 2013; Pascual-Gómez, Lorenzo-Llamas y Monge-López, 2015).

Otro término asociado a estas prácticas de evaluación en las que intervienen los agentes protagonistas del proceso de enseñanza y aprendizaje es la evaluación compartida, entendida como aquella en la que se realizan dinámicas dialógicas y

evaluativas sobre los procesos de aprendizaje (López-Pastor y Pérez-Pueyo, 2017). Pueden realizarse individualmente, en grupos de trabajo y en gran grupo. Tienen como objetivo acompañar y guiar los procesos de evaluación de una manera personalizada y transparente y/o realizar una calificación dialogada. Otros autores como Salinas (2002) han utilizado este concepto para referirse cuando los profesores de un centro comparten información sobre sus prácticas de evaluación y se coordinan para construir conjuntamente un sistema de evaluación que compartirán.

Por último, cabe destacar que, aunque la literatura científica ha demostrado la eficacia de la participación de los alumnos en la evaluación, a la práctica existen algunas resistencias y grandes temas que preocupan a los docentes y pueden frenar su aplicación. Según Taras (2015) las principales preocupaciones son si los alumnos también pueden poner nota y si la valoración que hacen es de calidad y debe o no ser considerada. Una respuesta inicial a estas cuestiones reside en tener presente que no podemos pedir que se evalúen sin enseñarles a hacerlo. Los alumnos deben aprender a evaluarse y, para ello, hay que integrar la autoevaluación como contenido de estudio (Brown y Harris, 2014), y planificar un proceso en el que se les enseñe a reconocer y enunciar cuáles son los pasos que deben tomar, los aspectos que deben considerar, cómo planificarse y qué estrategias les van a ayudar a hacerlo autónomamente (Sanmartí, 2010b). En el apartado sobre proceso de evaluación profundizaremos en ese aspecto.

Finalmente, debemos hablar de la evaluación externa, aquella realizada por agentes externos a los procesos de enseñanza y aprendizaje. Dentro de este grupo, las pruebas estandarizadas representan, en muchos contextos educativos, una de las principales actividades de evaluación por las que deben pasar los alumnos para dar cuenta de sus aprendizajes. En muchos países, como es el caso de España, las leyes educativas proponen una evaluación de este tipo en diferentes momentos de la escolarización obligatoria como estrategia para acreditar los resultados del sistema educativo.

Otros muchos países, participan en el *Program International Student Assessm* (a partir de ahora PISA en sus siglas en inglés), desarrollado por la Organización para

la Cooperación y Desarrollo Económico (OCDE), cuyo objetivo es comparar sistemas educativos a través del resultado de los aprendizajes, es decir, a través de una evaluación acreditativa. Una revisión de literatura de artículos sobre PISA señala que los estudios basados en los datos que ha aportado PISA han ayudado a desarrollar investigación educativa enfocada a mejorar la educación a raíz de los resultados negativo y, a la vez, que estas interpretaciones sobre los resultados hay que considerarlos con precaución (Hopfenbeck, Lenkeit, El Masri, Cantrell, Ryan y Baird, 2018).

Finalmente, aunque las pruebas estandarizadas provocan fuertes debates en la comunidad educativa y se deben analizar sus luces y sus sombras (Sanmartí y Sardà, 2007), sus presupuestos pedagógicos con relación a la evaluación de competencias, sí que pueden motivar a muchos docentes a transformar su evaluación, incluso a introducir modificaciones en su día a día de aula (Alcañiz y Cervera, 2014). Existen artículos que demuestran que el análisis y trabajo con los ítems de las pruebas PISA ayuda a los docentes a transformar su evaluación (Muñoz y Charro, 2017). De hecho, como veremos en los siguientes apartados del marco teórico, muchas de las características que son el núcleo de la evaluación de competencias están en consonancia con las bases psicoeducativas de las pruebas PISA. Este hecho también se sostiene con estudios que demuestran que la evaluación formativa produce mayores resultados en pruebas estandarizadas (William, Lee, Harrison y Black, 2009).

1.1.3. El proceso de evaluación.

Tal y como hemos desarrollado en los puntos anteriores, cuando la concreción de la evaluación cuente con su totalidad de componentes que la conforman dirigidos a mejorar los procesos de enseñanza y aprendizaje, estaremos hablando de una evaluación para enseñar y aprender mejor. El resto de las prácticas evaluativas, asociadas también al ámbito educativo, deberían ser, desde nuestra perspectiva, complementos que ayuden y no que interfieran y monopolicen el sentido de la evaluación, como ha venido pasando hasta el momento.

Pero no es suficiente con definir por qué, qué, cuándo, cómo, quién de la evaluación, también es necesario organizar los diferentes componentes en un proceso que nos permita implementar de forma eficaz el sistema de evaluación. Para ello, debemos determinar y planificar un proceso en diferentes fases y decisiones, que nos guiarán en su realización. Las fases del proceso evaluativo según Rodríguez (2000, citado en Mateo, 2000) son las siguientes:

1. Establecimiento de los objetivos de evaluación
2. Asignación de las tareas a realizar por el alumnado
3. Fijación de los criterios de realización de estas
4. Explicitación de los estándares o niveles de logro
5. Tomar las muestras de las ejecuciones de los alumnos
6. Valoración de dichas ejecuciones
7. Retroalimentar adecuadamente al alumno
8. Tomar decisiones

En este proceso, al contrario del procedimiento tradicional de evaluación, en el que se asigna una tarea de evaluación, se corrige y se retorna la calificación y los errores al alumno, se diseñan y comparten criterios de realización con los alumnos, se recogen diversidad de evidencias, se valoran dichas ejecuciones y se toman decisiones.

Otro modelo de diseño de fases para el proceso de evaluación es el de Jornet, González, Suárez y Peraless (2011), que destaca por centrarse en definir momentos que permitan realmente una evaluación de las competencias. En la Figura 1 puede observarse un esquema elaborado a partir del sistema de evaluación que proponen. Para ello, destaca aspectos que ayudan al docente a desarrollar las competencias clave en unidades más concretas con las que poder operar, como los criterios, los estándares y los indicadores, y acciones que permiten valorarlas, como situaciones, tareas, evidencias y un proceso de valoración de estas.

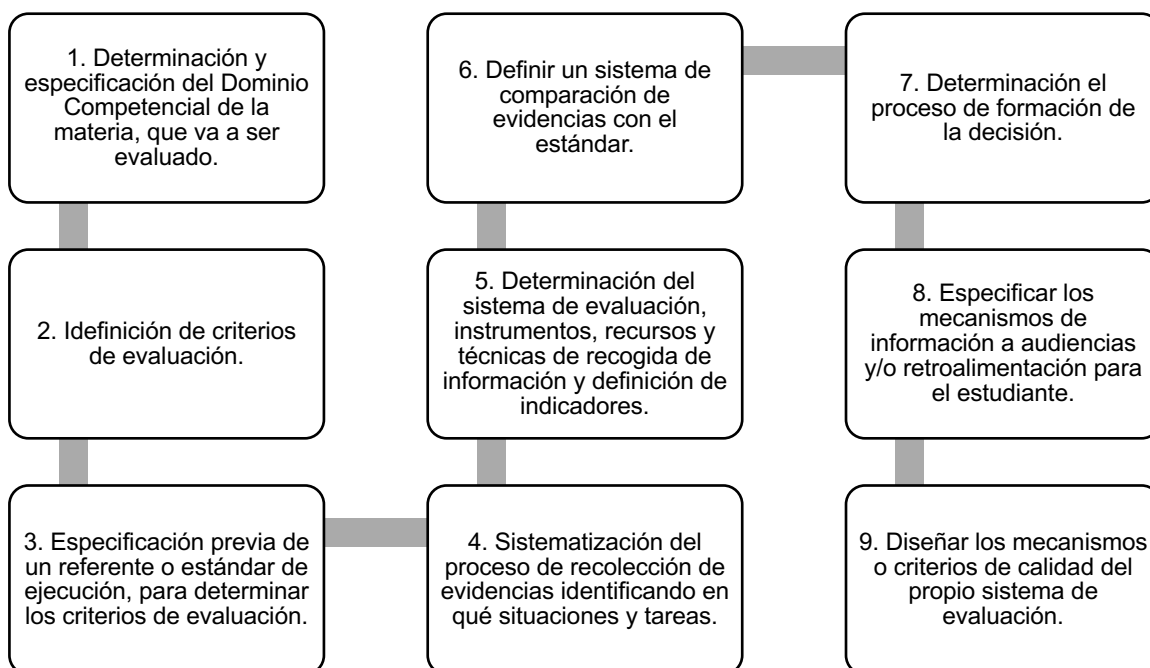


Figura 1. Sistema de evaluación de competencias según Jornet et al. (2011)

Obviamente, no todas las fases deben activarse cada vez que se inicia una nueva unidad didáctica o proyecto. El establecimiento de estándares y criterios puede realizarse una vez al curso y, de hecho, los currículums oficiales ya aportan información suficiente como para solo tener que decidir cuáles y en qué momentos. Por otro lado, el resto de las fases, deben iniciarse de forma recursiva para poder planificar la evaluación según el contexto y la situación de enseñanza y aprendizaje.

Tener presentes las fases y las concreciones para conseguir la evaluación de competencias es importante, un primer paso para avanzar hacia una evaluación formativa, sin embargo, muchas de estas fases pueden no aportar información suficiente al alumno o no lo hacen partícipe en este proceso. Es necesario ir más allá y hacer explícitos los momentos en los que la evaluación se integrará en los procesos de enseñanza y aprendizaje.

Al respecto, es interesante el modelo de Thomas y Cross (1993, citado en Mateo, 2000) en el que se realiza un ciclo de evaluación de aprendizaje en el aula, que consta de tres fases, planificación, implementación y respuesta, y diez pasos, que pueden observarse en la Tabla 6.

Tabla 6*Ciclo de evaluación de aprendizajes en el aula*

FASE	PASOS
Planificación	1. Elige el tema central 2. Céntrate en un objetivo o cuestión evaluable 3. Planifica un proyecto de evaluación en clase, centrándote en aquella cuestión u objetivo
Implementación	4. Explica la lección relacionada con ese objetivo o cuestión 5. Evalúa el aprendizaje del alumnado: recolectando información orientada a la retroalimentación 6. Analiza el proceso autorregulador del alumnado
Respuesta	7. Interpreta los resultados y formula una propuesta para mejorar los aprendizajes 8. Comunica los resultados y comprueba la respuesta 9. Evalúa los efectos de la evaluación sobre la enseñanza y el aprendizaje.
Nuevo ciclo	10. Diseña un proyecto de seguimiento de la investigación evaluativa en el aula

Nota. Extraído de Thomas y Cross (1993, p. 71)

Este ciclo de evaluación, respecto al proceso descrito anteriormente, aporta momentos en los que se evalúa de forma explícita cómo el alumno se autorregula. También, en la fase final de toma de decisiones, se contemplan pasos orientados a proponer mejoras y dejar que los alumnos las apliquen para volver a evaluar la efectividad de la enseñanza. Es posiblemente en este momento donde la evaluación se transforma en aprendizaje de forma más explícita.

Hasta ahora, los dos procesos descritos aportan información suficiente al docente para mejorar su enseñanza, en tanto que evalúa los aprendizajes y aportar retroalimentación para aprender y volver a evaluar los resultados. Sin embargo, todavía podríamos añadir algunos elementos más para que este proceso fuera una mejor oportunidad de aprendizaje para los alumnos. Al respecto, la propuesta de Sanmartí (2010) nos parece muy eficaz para desarrollar la evaluación desde su función reguladora. Se trata de añadir tareas que acompañen a los alumnos en su reflexión sobre cómo se aprende y no únicamente sobre lo aprendido. El objetivo es conseguir que los alumnos desarrollen su capacidad de autorregulación, a través de los siguientes momentos:

1. Identificar los objetivos de la actividad o conjunto de actividades.
2. Anticipar y planificar la acción necesaria para realizar la tarea.

3. Compartir los criterios de evaluación.

El primer paso es importante para que los alumnos, conociendo los objetivos y considerando las expectativas, puedan más adelante reconocer qué están trabajando y conectar con los propósitos que tenían, haciéndose preguntas relacionados con estos. Consiste en compartir con ellos el sentido de las actividades con el objetivo que después puedan comprobar si se han alcanzado; de esta manera, pueden implicarse activamente y podrán regularse de forma más efectiva. Para ello, es necesario explicitar cuáles son los objetivos y, aún más importante, regular la comprensión y las expectativas que se tienen sobre ellos. Se consigue a través de actividades en las que se hacen emerger y se organizan conocimientos previos y contenidos, a través de organizadores gráficos o cuestionarios, y en las que se consensuan los objetivos, a través de actividades como debates o contratos.

El segundo paso es necesario para que los alumnos puedan organizarse y monitorear las actividades que tienen que realizar. Se trata de que los alumnos concreten el proceso de resolución definiendo cuáles son los pasos que un profesional llevaría a cabo para conseguir una resolución óptima. Este proceso, recogido en un documento, les servirán posteriormente para poder evaluar si están realizando correctamente la actividad. Cualquier instrumento que permita organizar y reflexionar sobre el pensamiento, como las guías metacognitivas, los organizadores gráficos de destrezas de pensamiento o los diarios de clase, y/o que permita organizar los pasos de resolución de una tarea, como las bases de orientación o los esquemas o mapas conceptuales, pueden responder a este objetivo. Es especialmente relevante que los alumnos participen en el diseño de estos instrumentos para poder interiorizar los pasos, como también lo es que estos se compruebe la utilidad de estos y se modifiquen si es preciso.

El tercer paso, es imprescindible para que los alumnos puedan saber cuáles son los aspectos relevantes que tienen que considerar y cómo deben enfocarlos para conseguir una resolución de calidad. Es decir, hasta ahora hemos indicado en qué nos centraremos y cómo llevarlo a cabo, pero no qué tendremos que hacer específicamente para resolver la tarea con éxito y calidad. Para ello, es necesario

que los alumnos identifiquen y consensuen los criterios de realización y de resultado, los cuáles indican los aspectos importantes y en qué grado pueden desarrollarse, respectivamente. Para consensuar estos criterios, pueden realizarse tareas como el análisis de trabajos previos y debates para poner en común los resultados, siempre con la guía del docente, y construir conjuntamente instrumentos como rúbricas de evaluación o tablas de indicadores. Posteriormente, como en el paso anterior, es necesario analizar la efectividad de estos utilizándolos, por ejemplo, en actividades de autoevaluación y coevaluación.

Una vez definidas cuáles son las fases y actividades específicas para conseguir un proceso de evaluación realmente formativo que fomente la autorregulación de los aprendizajes, por último, nos parece importante desarrollar un aspecto de vital importancia para la calidad de esta evaluación formativa: la retroalimentación que realizan los docentes.

1.1.3.1. La retroalimentación durante la evaluación.

La retroalimentación durante la evaluación cumple varias funciones: ayudar a los alumnos en la realización de las actividades, guiar la autorregulación de los aprendizajes y construir una percepción positiva sobre la evaluación. Consiste, en su definición más acotada, en reducir la distancia entre aquello que realiza o piensa el alumno y el resultado esperado. Puede llevarse a cabo durante todo el proceso de evaluación y con una frecuencia indeterminada, en tanto que es dependiente de las necesidades del alumnado y de las habilidades del docente (Voerman, Meijer, Korthagen y Simons, 2012).

La retroalimentación no debe confundirse con las prácticas más tradicionales en las que el profesor devolvía los resultados indicando lo correcto o lo incorrecto, sino que es un aspecto clave de la evaluación formativa. Así pues, va más allá, y tiene como objetivo ayudar a los alumnos a generar mecanismos cognitivos que les permitan redirigir su pensamiento y reconstruir sus representaciones sobre el aprendizaje (García-Jimenez, 2015). Y, para realizar esto, es imprescindible tener claros los criterios de evaluación, haber analizado evidencias que muestren el desempeño y conocer el perfil de aprendizaje del alumno.

En la literatura sobre retroalimentación encontramos dos términos especialmente relevantes: el *feedback* y el *feedforward*. Cuando hablamos de *feedback*, nos referimos a la retroalimentación centrada en aportar información sobre la tarea realizada en la situación de aprendizaje que se está desarrollando, por su parte, el *feedforward* se dirige a aportar información útil para realizar mejor la tarea en futuras situaciones similares. La gran diferencia es esta perspectiva en el comentario que se realiza con el *feedforward*, en la que se informa también de los motivos que han podido causar un tipo de resolución y como superarlos en un futuro (Ion, Silva y Cano, 2013).

Por último, el impacto de la retroalimentación ha quedado demostrado en varias investigaciones y va desde la valoración positiva por parte de los estudiantes en tanto que declaran que les ayuda a regular los aprendizajes, hasta la aparición de cambios positivos en los aspectos emocionales y personales de los participantes (Canabal y Margalef, 2017; Tillema, Leenknecht y Segers, 2011). Además, la retroalimentación también aporta información al docente mediante la interacción y escucha de los alumnos, lo cual favorece el ajuste de la enseñanza a los diferentes perfiles (Wiggins, 2011).

En conclusión, en el diseño de un proceso de evaluación se deben planificar tanto actividades que permitan recoger una gran variedad de evidencias sobre la ejecución y desempeño de los alumnos en diferentes momentos después de la enseñanza y la retroalimentación, como actividades que permitan al alumnado conocer los objetivos, el proceso y los criterios de calidad mediante los que pueda regular y mejorar su aprendizaje. Las primeras responden más a una evaluación de los resultados enfocada a mejorar la enseñanza y las segundas a una evaluación de los resultados enfocada a mejorar el aprendizaje con la participación de los alumnos. Y, en este último grupo, es de vital importancia la comunicación que tiene el profesor durante la evaluación, la cuál debe estar dirigida a ofrecer retroalimentación constructiva y que ayude a sus alumnos a representarse el error como una oportunidad de aprendizaje.

Para cerrar este apartado sobre la evaluación en y de los procesos de enseñanza y aprendizaje, en el cuál hemos descrito cuáles son los orígenes de este tipo de evaluación, cuáles los componentes a considerar para definirla y cuál el proceso para llevarla a cabo, queremos señalar que esta es una de las tantas posibilidades, concreciones y cambios educativos que nos puede ofrecer el debate sobre la evaluación educativa. De hecho, como en algún punto hemos mencionado, los resultados de carácter acreditativo nos pueden servir para iniciar una evaluación del profesorado y del propio sistema educativo. En este sentido, son muchos los temas que no hemos tratado a causa de los objetivos de esta investigación, como son: la participación y comunicación de los resultados a las familias, la evaluación del profesorado, la cultura de evaluación de centro, etc.

Sin embargo, nos centramos en profundizar en la evaluación de los procesos de enseñanza y aprendizaje de forma premeditada, conscientes de que "el incremento de la consciencia en los procesos de aprendizaje constituye uno de los más importantes retos educativos a los que la evaluación puede proporcionar soluciones" (Mateo, 2000, p. 17). Añadiríamos que es una de sus responsabilidades para una enseñanza de calidad. También porque, como hemos indicado en varias ocasiones, la evaluación debe estar en consonancia con los nuevos principios y teorías sobre el aprendizaje, dentro del marco de la escuela inclusiva y de la enseñanza competencial hacia el cual se están dirigiendo las escuelas, por lo menos en el contexto español.

Por ello, hemos decidido enfocar nuestro análisis a una evaluación alternativa, concretamente en la evaluación auténtica, con el fin de defender sus potencialidades dentro de este objetivo. En el siguiente apartado vamos a desarrollar cuáles son las bases, principios y características de este tipo de evaluación, señalando cómo da respuesta a las actuales necesidades de la evaluación de y en los procesos de enseñanza y aprendizaje en el contexto actual educativo.

1.2. La Evaluación Auténtica

La amplia gama de opciones a escoger en el momento de diseñar un sistema de evaluación puede aturdir al maestro y resultarle confusa, especialmente cuando no se tiene claro el propósito central de la evaluación en el sistema educativo. Nos referimos a la función de estar al servicio del aprendizaje de los alumnos, los cuáles son los verdaderos protagonistas. Y esto sucede dentro de un sistema de enseñanza que en la actualidad está orientado a las competencias y a la inclusión de todos los alumnos en el sistema educativo.

Este enfoque actual hacia el que se dirige la construcción el sistema educativo determina claramente la toma de decisiones sobre para qué, qué y de qué manera se evalúa. Así pues, aunque hay algunas opciones sobre cómo evaluar que pueden combinarse entre ellas, como por ejemplo emplear pruebas tipo test y a la vez carpetas de aprendizaje; también existen otras que no deberían ser opciones únicas, como por ejemplo evaluar solo mediante pruebas tipo test al final de la unidad didáctica, y otras opciones que no se pueden obviar, como es el caso de no posibilitar que el alumno se autoevalúe, especialmente si queremos que éste sea el principal actor de su aprendizaje.

Consecuentemente, en el momento de escoger las características que definen un sistema de evaluación, debemos elegir aquellas que estén en sintonía con el enfoque de enseñanza competencial dirigido a un aprendizaje activo y significativo. De las diferentes tipologías de evaluación existentes, en esta memoria defenderemos el enfoque la evaluación auténtica pues consideramos que es el que mejor responde a una enseñanza basada en competencias.

La evaluación auténtica nació como una contrapropuesta a la evaluación tradicional, a la par que el resto de los modelos alternativos de evaluación. Está vinculada, por tanto, a la evaluación formativa, la evaluación basada en problemas y la evaluación del desempeño, entre otras. En su intento de alejarse de una evaluación sumativa que pretenda clasificar a los alumnos comparando sus resultados mediante el uso de pruebas escritas, la evaluación auténtica está al

servicio de los aprendizajes de los alumnos mediante el uso de actividades prácticas y formativas. Como señalan Condemarín y Medina (2000):

La fuerza de la perspectiva de evaluación auténtica reside en su concepción de la evaluación como parte integral y natural del aprendizaje, y en las oportunidades que ofrece para utilizar las reales actividades cotidianas que ocurren en la sala de clases como fuente de recolección de información y base para la toma de decisiones. (p.15)

Además, como alternativa a las tareas de evaluación más predominantes, basadas en ejercicios que piden reproducir información, la evaluación auténtica propone actividades que reclaman la resolución de problemas complejos. Siguiendo a Monereo (2009), la evaluación auténtica se caracteriza:

[...] por valorar especialmente el proceso de decisión necesario para resolver un problema complejo, en el que deban activarse y aplicarse coordinadamente distintos conocimientos y competencias, y haya que demostrar la capacidad estratégica imprescindible para autorregular la propia conducta y ajustarse a los cambios inesperados, justificando posteriormente las acciones efectuadas. (p.18)

Pero si hay una característica de la evaluación auténtica que de verdad la hace destacar por encima de las evaluaciones coetáneas como las mencionadas evaluación formativa y evaluación del desempeño, es su conexión con el mundo real. Es decir, que el problema o la situación a resolver mediante la que el alumno demostrará su aprendizaje y podrá continuar aprendiendo, se caracteriza por ser realista y relevante para este.

La similitud entre lo que se evalúa y las situaciones reales que el estudiante podría afrontar es un elemento que los autores de referencia han señalado como indispensable. Partimos pues de la idea de que la autenticidad guarda un alto grado de conexión con la "realidad", esto es, con el contexto y las situaciones problemáticas que le son propias, una relación que se hace presente tanto en las condiciones de ejecución como en las demandas cognitivas exigidas. Y, como la realidad es compleja en sí misma, también se caracteriza por demandar una

resolución estratégica, tareas colaborativas, el uso de los recursos disponibles y la reflexión sobre el proceso seguido, con el fin de mejorarlo en futuras situaciones.

En este apartado vamos a desarrollar en profundidad estas características en base a la literatura publicada al respecto. También explicaremos cuáles son las bases psicoeducativas y principios que la conforman. Y, para acabar, plantearemos de qué manera puede ponerse en práctica una evaluación auténtica, descubriendo algunas experiencias que han resultado exitosas en su implementación, y cuáles son las decisiones y barreras que se deben considerar.

1.2.1. Bases psicoeducativas, principios y características.

La literatura sobre evaluación auténtica se ha visto en auge considerablemente durante los últimos años. Cada vez son más los estudios en los que se conceptualiza el término y se diseñan sistemas y prácticas de evaluación auténticas que, a la vez, den respuesta a las nuevas tendencias en pedagogía. En consecuencia, estos estudios comparten unas bases psicoeducativas comunes a partir de las cuáles se definen los principios y características que constituyen la evaluación auténtica.

En primer lugar, con relación a las bases psicopedagógicas, encontramos que las experiencias de evaluación auténtica se fundamentan en el aprendizaje significativo y en el enfoque de enseñanza competencial. Conecta con las perspectivas sociocognitivas y constructivistas, en las que se considera que el aprendizaje sucede de forma situacional y contextualizada y que parte de los esquemas cognitivos previos de los alumnos a través de la actividad. En palabras de Álvarez (2005):

La práctica educativa, con foco hacia el aprendizaje y desarrollo de competencias, requiere concebir el aprendizaje como una construcción activa y social de conocimientos, de forma reflexiva y, sobre todo, con estrecho vínculo al contexto de referencia del objeto de estudio. (p.52)

Consecuentemente, si el aprendizaje debe ser experiencial, contextualizado y progresivamente más complejo, la evaluación debe también serlo, y la evaluación

auténtica responde plenamente a esos requisitos gracias a su esfuerzo por conectar con el "mundo real", proponiendo actividades que permiten poner en práctica, en contextos reales o de realidad simulada, lo aprendido. Además, de esta manera, la evaluación auténtica da respuesta al problema de la significatividad del aprendizaje, entendiendo que la transferencia de lo aprendido a un contexto real no se produce de forma automática ni directa y, por tanto, debe aprenderse de forma contextualizada. Es decir, si sabemos para qué sirve un conocimiento, pero no lo usamos ni desarrollamos, es muy probable que cuando nos encontremos en una situación donde lo necesitemos, no sepamos activarlo ni utilizarlo eficazmente y, por lo tanto, no lleguemos a ser competentes.

Pero, aunque estas bases psicoeducativas son comunes a la mayoría de las propuestas evaluativas centradas en la autenticidad, las características específicas que constituyen un sistema de evaluación auténtica no son siempre las mismas. A raíz de la diversidad y número de experiencias reales en el aula ha emergido una gran variedad de prácticas y características que la definen. Algunas de estas características son coincidentes con otras tipologías de evaluación, como avanzábamos anteriormente, y otras son propias de la evaluación auténtica, aunque todas, desde nuestra perspectiva, y como mostraremos en los siguientes puntos de este apartado, están interconectadas.

De esta manera llegamos a la definición de un conjunto de principios de evaluación auténtica. Aunque más de un autor enuncia estos principios en sus publicaciones, hemos seleccionado los de Vallejo y Molina (2014) por ser una muestra representativa de la gran variedad de nociones que puede llegar a englobar una evaluación auténtica. Estos principios son los siguientes:

- a) Explorar los aprendizajes que requieren habilidades cognitivas y acciones complejas, no el simple recuerdo de información o la ejercitación rutinaria; b) Seleccionar o desarrollar tareas auténticas que representen tanto el contenido como las habilidades centrales –en términos de los aprendizajes más importantes– y, de esta manera, conjugar la enseñanza con la evaluación; c) Proporcionar a los alumnos los apoyos necesarios para que comprendan y realicen la actividad, así como para entender las expectativas

existentes en función del nivel de logro esperado; d) Comunicar con claridad las expectativas de ejecución en términos de criterios consensuados con el grupo, mediante los cuales se juzgará dicha ejecución, al tiempo que se deben generar las condiciones y mecanismos necesarios que permitan registrar el avance de los alumnos; e) Incluir espacios de reflexión en torno a los aprendizajes logrados, a la enseñanza que los posibilitó y a los mecanismos de evaluación que se emplearon, recuperando posteriormente dichas reflexiones como elementos de realimentación y propuestas para la mejora. (Vallejo y Molina, 2014, p.18)

Si analizamos estos principios, encontraremos categorías como plantear la evaluación como un reto, atender la diversidad de los alumnos o posibilitar que la evaluación sea una oportunidad de aprendizaje. Además, esto puede traducirse en una amplia variedad de prácticas e instrumentos de evaluación como el análisis de casos, la coevaluación y la autoevaluación, el uso de rúbricas, realización de las carpetas de aprendizaje o la resolución in situ de problemas.

Llegados a este punto, creemos que es necesario conocer cuáles son genuinamente las características que la literatura sobre evaluación auténtica señala como las más representativas de este enfoque. Para ello vamos a considerar dos artículos de revisión de la literatura sobre evaluación auténtica. El primer estudio de revisión de literatura sobre evaluación auténtica es el realizado por Frey, Schmitt y Allen (2012) en el que se parte del objetivo de aclarar qué significa realmente la palabra auténtica dentro del concepto, puesto que en algunos estudios se utiliza como sinónimo de "fidelidad" con el proceso de enseñanza y aprendizaje y en otros como sinónimo de "conexión con el mundo real". Consiste en el análisis de 109 textos seleccionados de una muestra más amplia, considerando su aportación al concepto de evaluación auténtica. En los resultados se identifica nueve características relativas a la "autenticidad", las cuales se agrupan en tres grandes categorías según se refieran al contexto de evaluación, al rol del alumno o a la puntuación de la actividad evaluada (ver Tabla 7).

Tabla 7*Dimensiones y características de la evaluación auténtica según Frey et al. (2012)*

El contexto de evaluación	El rol del estudiante	La puntuación
<ul style="list-style-type: none"> • Actividad o contexto realista • La tarea se basa en el rendimiento • La tarea es cognitivamente compleja 	<ul style="list-style-type: none"> • Se requiere la defensa de la resolución o del producto • La evaluación es formativa • Los estudiantes colaboran entre ellos o con el maestro 	<ul style="list-style-type: none"> • Los criterios de calificación son conocidos o desarrollados por los estudiantes • Se utilizan múltiples indicadores o evidencias para puntuar • El dominio es el rendimiento esperado

Nota. Extraído de Frey et al. (2012)

Estas nueve características no están presentes de la misma manera en todos los artículos. Los resultados del estudio señalan que, aunque comúnmente en la evaluación auténtica se asume como característica principal el realismo de las tareas o de la propuesta en sí, con cierta frecuencia ni siquiera se menciona. También, en el estudio se realiza una clasificación de los artículos según los niveles de enseñanza en los que se lleva a cabo la experiencia, para poder explorar en qué contextos educativos logra más aceptación, siendo la educación primaria y secundaria obligatoria en las que se ha implementado más la evaluación auténtica.

En el segundo estudio, de Villarroel, Bloxham, D. Bruna, C. Bruna y Herrera-Seda (2017), se realizó una revisión de 125 artículos sobre evaluación auténtica publicados entre 1988 y 2017. Como conclusión se identifican tres dimensiones que están reflejadas en todos los estudios: dimensión de realismo, dimensión de reto cognitivo y dimensión de juicio de evaluación, junto con sus características (ver tabla 8).

Tabla 8*Dimensiones y características de la evaluación auténtica según Villarroel et al. (2017)*

Realismo	Reto cognitivo	Juicio de evaluación
<ul style="list-style-type: none"> • Tareas similares al mundo real / laboral. • Valor práctico. • Problemas contextualizados en la vida cotidiana. • Relevancia más allá del aula. • Rendimiento auténtico. • Competencias para el desempeño laboral. 	<ul style="list-style-type: none"> • Pensamiento de orden superior. • Habilidades para resolver problemas. • Habilidades para tomar decisiones. 	<ul style="list-style-type: none"> • Retroalimentación. • Sentido formativo. • Criterios de evaluación conocidos a priori.

Nota. Adaptado de "Dimensions of the authentic assessment (AA)" en Villarroel et al., 2017, p. 4)

Según Villarroel y sus colegas (2017) la mayor parte de los estudios analizados, concretamente un 71%, destacan como característica definitoria aspectos relativos a la autenticidad -dimensión de realismo-; aproximadamente la mitad, un 55%, a características relacionadas con tareas y problemas complejos -dimensión reto cognitivo- y en menor medida, un 38%, a aspectos relacionados con emitir y compartir el análisis de las evidencias y resultados -dimensión juicio de evaluación-. A partir de esta identificación de características básicas, los autores proponen un método para ayudar a los docentes a implementar una evaluación auténtica en cualquier contexto educativo, incluso en aquellos centros que mantienen una fuerte tradición en evaluación basada únicamente en la aplicación de pruebas. Aunque reconocen que es difícil, sostienen que el trabajo sobre las tres dimensiones y el modelo basado en diferentes etapas posibilita que cada docente pueda empezar por realizar pequeños cambios en sus prácticas e ir avanzando a su ritmo, lo cual culminaría en una transformación de sus pruebas.

En síntesis, la literatura nos muestra que hay una gran variedad de experiencias educativas sobre evaluación auténtica en las que se enfatizan unas características u otras según el contexto y el propósito de la situación evaluativa. Podemos afirmar que la evaluación auténtica es compleja en tanto que considera muchos aspectos y acoge en su definición prácticas de otras corrientes evaluativas. Pero, después del análisis, también concluimos que hay tres dimensiones que son comunes y

representativas de la evaluación auténtica: una, su propósito competencial en tanto que existe una contextualización -caracterizada por el realismo, la relevancia, la practicidad, etc.- y una complejidad cognitiva elevada -caracterizada por suponer un reto cognitivo, promover el desarrollo de habilidades, basarse en el rendimiento, etc.-; dos, su objetivo de que suponga una experiencia de aprendizaje -apoyada por el hecho de compartir criterios evaluativos, indicadores, ofrecer retroalimentación constructiva, etc.-; tres, su carácter inclusivo en tanto que atiende la diversidad de los alumnos -caracterizado por proponer tareas de colaboración, ofrecer ayudas y recursos, etc.-.

Estas tres dimensiones se corresponden además con los temas claves en el sistema educativo actual que hemos señalado anteriormente: el enfoque de enseñanza por competencias, la evaluación formativa y la atención a la diversidad. Así pues, en el siguiente apartado se señalarán algunos de los aspectos de la evaluación auténtica que contribuyen a dar respuesta a estos retos educativos. Además, defendemos que para llevar a cabo una evaluación auténtica es necesario evaluar competencialmente, promover el aprendizaje y hacerlo de forma inclusiva, desglosando las dimensiones en sus características principales y destacando cómo se interrelacionan entre sí para dotar de sentido unitario a un sistema de evaluación auténtica. Posteriormente, analizaremos algunas prácticas de evaluación auténtica que se han demostrado especialmente eficaces.

1.2.1.1. La autenticidad requiere evaluar competencialmente.

La evaluación auténtica se caracteriza por la observación y valoración de la actuación competente del alumno dentro de situaciones realistas y relevantes para este. Si consideramos que ser competente significa utilizar aquellos saberes necesarios para resolver una situación de forma eficaz en un contexto determinado, para evaluar el desarrollo competencial de los alumnos tenemos que diseñar actividades complejas donde podamos observar cómo estos movilizan conocimientos y recursos para responder al reto que se les propone. Por consiguiente, es necesario que las actividades de evaluación, en mayor o menor grado, sean auténticas, contemplen una situación contextualizada que suponga un

reto, se desarrollen a través de tareas graduadas según su complejidad cognitiva y de condiciones de resolución parejas al contexto recreado (ver Tabla 9).

Tabla 9

Descripción de las características de una evaluación auténtica que evalúa competencias

Característica	Descripción
Situación problemática contextualizada	Consiste en describir una situación en la que se presente un problema que pueda encontrarse en un contexto profesional, personal o social, detallando todas las características que la describen (quién participa, dónde sucede, tiempo y recursos...)
Actividades auténticas	Se trata de que las tareas o acciones a realizar para participar en la situación o resolver el problema sean realistas, por su similitud con el mundo real, relevantes en tanto que significativas para los alumnos y socializadoras porque permiten poner en práctica el rol propio del contexto (habilidades sociales, discurso característico, equipo, materiales, instrumental específico, etc.)
Complejidad cognitiva	Consiste en que el proceso de participación o resolución del problema implique en la superación de demandas de baja, media y alta complejidad cognitiva, que se sucedan de forma progresiva de menos a más dificultad.
Recursos y condiciones de resolución auténticas	Se trata de recrear, al máximo posible, la situación problemática contextualizada, ajustando las condiciones como el tiempo, los materiales disponibles, las situaciones de interacción o el espacio físico en el que ocurra, etc.

Estas características están conectadas entre sí de forma que la concreción de una de ellas provoca el ajuste del resto. Por ejemplo, entender la competencia como la resolución eficaz de problemas complejos de forma situada nos conduce a definir situaciones contextualizadas donde se puedan poner en práctica dicha resolución. Bajo el objetivo de formar ciudadanos competentes, debemos identificar los contextos y prácticas socioculturales en las que participarán y las condiciones, elementos y tareas que deberían llevar a cabo para resolver los problemas que puedan generarse. Por lo tanto, uno de los pasos claves para evaluar competencias es situarse en un contexto profesional, social o personal e identificar cuáles son los problemas prototípicos o emergentes que le son propios (Ashford-Rowe, Herrington y Brown 2014; Monereo, 2009).

Con el planteamiento de una situación contextualizada es casi inevitable que para definirla partamos del mundo real y, con ello, que las actividades que planteemos sean auténticas, en tanto que integran conocimientos y habilidades tal y como se

utilizarían en la realidad (Darling-Hammond y Snyder, 2000). Y, como la realidad no es lineal ni estática, sino compleja y dinámica, al diseñar actividades auténticas, aparecen tareas o acciones que son cognitivamente complejas, es decir, que permiten una gama amplia de respuestas y la construcción de un proceso de resolución concreto (A. Herrington y J. Herrington, 2006). Así pues, las actividades auténticas que parten de la situación contextualizada nos conducen a la formulación de un reto cognitivo, que se resuelve con diversidad de acciones ordenadas de menor a mayor complejidad cognitiva (Avery, Freeman y Carmichael, 2012).

Todo ello, situación contextualizada, autenticidad y complejidad, responden pues a la lógica de un diseño competencial. Pero para poder evaluar el desempeño real es necesario un paso más: observar la puesta en acción de los alumnos. Así pues, las actividades o tareas propias de la resolución del problema deben estar en sintonía con la autenticidad de la situación contextualizada, de hecho, todos los elementos que componen la evaluación deben ser auténticos (Gulikers et al., 2004). Específicamente nos referimos al grado en que las tareas que se proponen son realistas, es decir, tienen una similitud con actividades homólogas en la vida real; son relevantes, en tanto que existe una utilidad explícita para su uso en contextos cercanos al estudiante; y son socializadoras, puesto que implican posicionarse en un rol específico de la comunidad en la que se sitúa la práctica.

Estos tres aspectos, realismo, relevancia y socialización deben, por supuesto, guardar una correspondencia clara con el problema que se plantea y adaptarse al nivel en el que se desarrolla la enseñanza ajustando las condiciones físico-sociales, la aplicabilidad y los conocimientos previos necesarios. También se ha de considerar el contexto escolar en el que se proponen y ajustarse a las posibilidades de implementación habituales de la práctica docente. Lo que nos lleva a hablar de la última característica: condiciones de resolución acordes con la situación auténtica (A. Herrington y J. Herrington, 2006). Si queremos observar el desempeño de los alumnos tal y como sería en la realidad, debemos ajustar al máximo posible las condiciones de resolución tales como los recursos con los que se cuenta, el tiempo de resolución, el contexto físico, etc.

En síntesis, una evaluación auténtica requiere la construcción de una situación auténtica y compleja que presente un reto cognitivo a los alumnos, a través del planteamiento de un problema y de tareas realistas, relevantes y socializadoras con sus condiciones de resolución ajustadas al contexto. Este diseño posibilita la evaluación de las competencias del alumno y la transferencia de los aprendizajes a la práctica.

No obstante, evaluar competencias es únicamente uno de los requisitos necesarios para adaptarse a una evaluación que dé respuesta a los retos de la evaluación en el contexto educativo actual. Como desarrollamos en el anterior apartado, es necesario que la evaluación promueva el aprendizaje. De hecho, en la realidad, los ciudadanos tanto en sus vidas profesionales como personales aprenden mientras desarrollan sus competencias. Por tanto, la autenticidad requiere que la evaluación también sea formativa, como veremos a continuación.

1.2.1.2. La autenticidad requiere que la evaluación sea formativa.

La evaluación auténtica se caracteriza por ofrecer oportunidades para que los alumnos puedan aprender durante y después de la evaluación. Durante, descubriendo nuevos contextos y problemas en los que poder aplicar sus conocimientos y habilidades para entrenar sus competencias. Después, reflexionando sobre su actuación y recibiendo retroalimentación que les ayude a mejorar y continuar aprendiendo.

Con todo, para posibilitar el aprendizaje gracias a la evaluación, es necesario que los alumnos entiendan que el error es una oportunidad de aprendizaje y que, por lo tanto, deben descubrir cómo mejorar reflexionando sobre su proceso, apoyándose en los criterios de calidad compartidos durante la evaluación y analizando las recomendaciones recibidas de una persona más experta (ver Tabla 10). En el momento en el que los alumnos se convierten en parte del proceso de evaluación, aumenta su implicación y contribuye al desarrollo de su capacidad de aprender a aprender en tanto que se les enseña a anticipar, supervisar y mejorar la calidad de sus trabajos (Kearney, 2013; Kearney, Perkins y Kennedy-Clark, 2016).

Tabla 10*Descripción de las características de una evaluación auténtica orientada al aprendizaje*

Característica	Descripción
Comunicación constructiva de los resultados	Consiste en realizar una retroalimentación de los resultados que permita aprender cómo hacerlo mejor y no solo cuáles son los errores, para que el alumno entienda que el error es una oportunidad de aprendizaje.
Evaluación del proceso y recolección de evidencias	Se trata de recoger evidencias (producciones, lista de indicadores de progreso, registro de observaciones, etc.) que den cuenta de la resolución estratégica del alumno (antes, durante y después) y permitirle que reflexione sobre su proceso comparando evidencias iniciales y finales.
Criterios de calidad compartidos	Consiste en determinar conjuntamente o compartir (antes, durante y/o después) con los alumnos cuales son los criterios para unos resultados de calidad, mediante el uso de instrumentos (bases de orientación, rúbricas, tablas metacognitivas, etc.) que faciliten la comprensión por parte del alumno de aquello que puede mejorar para conseguir mayores resultados.
Retroalimentación	Se trata de posibilitar tareas que aporte retroalimentación sobre los procesos y progresos de los alumnos, sea mediante una autoevaluación, coevaluación o heteroevaluación, mediante la que se explique cuáles son los cambios que puede realizar para corregirse y mejorar.

Estas características también están ligadas entre sí de forma que la concreción de cada una exige la concreción de otra. Una de las características básicas de la evaluación orientada al aprendizaje, es conseguir que los alumnos entiendan que el error es una oportunidad de aprendizaje. Se trata, pues, de que la evaluación no sea una actividad final, sino una actividad más de aprendizaje (Heritage, 2018). Esto se consigue mediante una comunicación constructiva de los resultados donde las calificaciones numéricas no sean el centro de la retroalimentación y ofreciendo la posibilidad de repetir la evaluación. Pero, para conseguir que los alumnos comprendan que la evaluación sirve para mejorar, es necesario también ofrecerles la oportunidad de reflexionar sobre su propio proceso y resultados. Por ello, es necesario que el docente programe momentos durante la evaluación en los que los alumnos deban relatar el proceso que están siguiendo, antes, durante y/o después de las tareas o la actividad, y tengan que recoger diferentes evidencias para evaluarlas después (Mc.Laren, 2012).

De la necesidad de permitir momentos de reflexión sobre el proceso y progreso, es fácil deducir que los alumnos necesitan ayuda para realizar esta reflexión, es decir,

que debemos proporcionarles herramientas para que sepan qué aspectos deben atender y cómo mejorarlos. Para ello, determinar y compartir los criterios e indicadores de evaluación ha demostrado ser un paso decisivo (García-Jiménez, 2015). Se trata de, en la medida de lo posible, establecer juntamente con los alumnos cuáles son los criterios de evaluación, así como proporcionarles instrumentos como las bases de orientación, las rúbricas, las pautas metacognitivas, etc., que puedan utilizarse durante la evaluación para conseguir mejores resultados (Dawson, 2015). Esto conecta con la dimensión de competencialidad, en tanto que en los contextos reales es habitual que los profesionales de un área tengan un manual o guía para la realización de sus tareas y tomen diferentes momentos en su creación para revisar y mejorar sus productos.

Por último, si necesitan ayuda para reflexionar sobre su proceso y mejorar, es necesario, además de ofrecerles los criterios, enseñarles a autoevaluarse con estos criterios. Actualmente existen muchas y variadas estrategias pedagógicas para su implementación (Panadero y Alonso-Tapia, 2013). Una estrategia para conseguirlo es que los docentes, en su rol de guías de aprendizaje, realicen una retroalimentación constructiva que, además, les sirve de modelo a los alumnos. Consiste en valorar tanto el proceso como los resultados y ofrecerles orientaciones que les indiquen cómo mejorar (Hill y West, 2019). No se trata tanto de decirles dónde está el error, sino de acompañarlos para que lo descubran y aprendan cómo rectificarlo. Esta premisa nace de la idea de que el alumno puede corregir sus errores porque es el que mejor puede descubrir cómo sucedió y qué hacer al respecto (Sanmartí, 2007). Otra estrategia, consiste en proponer actividades de coevaluación en las que sean los iguales quienes valoran la realización de las actividades. Esta acción tiene beneficios en ambos alumnos, puesto que no solo el evaluado recibe ayuda, sino que el evaluador aprende en tanto que interioriza los criterios de evaluación a través de su aplicación.

En síntesis, la evaluación auténtica fomenta que los alumnos aprendan a autorregularse como acción estratégica para la resolución eficaz de situaciones contextualizadas que presenten un reto real (Álvarez, 2009). Con ello, pretenden ir más allá de la evaluación de competencias y servir a la vez como proceso formativo en tanto que la resolución ayuda a los alumnos a aprender a prepararse para su

vida profesional. Además, los estudios señalan que los procesos de evaluación formativa mejoran la motivación (Meusen-Beekman, Joosten-ten Brinke, & Boshuizen, 2016).

Sin embargo, para conseguirlo es necesario considerar que no todas las personas son iguales y, por tanto, los alumnos tienen experiencias previas distintas, unas habilidades desarrolladas de diferente manera y unas preferencias de aprendizaje que hacen que se enfrenten de manera distinta a los problemas. Tener presente esta diversidad es un acto de equidad y permite potenciar el aprendizaje y el desarrollo de competencias como veremos en el siguiente punto.

1.2.1.3. La autenticidad requiere evaluar de forma inclusiva.

La evaluación auténtica se caracteriza por respetar y favorecer las diferentes habilidades y formas de aprender que tienen los alumnos. Los contextos reales, tanto personales, sociales como profesionales, están constituidos por personas muy diversas que actúan según sus capacidades, habilidades e intereses, de forma que un mismo problema puede resolverse y comunicarse de muchas maneras según sean las personas que participen. Por consiguiente, una evaluación auténtica es inclusiva en tanto que ofrece diversidad de propuestas para recibir, participar y comunicar los resultados de las actividades, diferentes ayudas según las necesidades de los alumnos y elementos que promueven su participación e implicación (Tabla 11).

Tabla 11

Descripción de las características de una evaluación auténtica que atiende la diversidad

Característica	Descripción
Diversidad de recursos, actividades y formas de comunicar	Consiste en diversificar tanto los <i>inputs</i> y <i>outputs</i> de los que consta la actividad evaluativa (formato escrito, oral, audiovisual, auditivo...) como los recursos que se ofrecen de forma que los alumnos tengan la posibilidad de recibir, analizar y comunicar la información de diferentes formas según su perfil de aprendizaje.
Condiciones ajustadas a los perfiles de aprendizaje	Se trata de ofrecer a los alumnos ayudas y recursos (guías, más actividades colaborativas, acceso a contenidos...) para que puedan participar en la resolución de la actividad de evaluación sin rebajar la complejidad de esta.
Ajuste a los intereses y motivaciones de los alumnos	Consiste en introducir elementos (temáticos, acceso y comunicación de la información, dinámicas de acción, etc.)

que estén en sintonía con los intereses de los alumnos e incrementen así su participación y motivación durante la evaluación.

También estas características están relacionadas entre sí en tanto que comparten el mismo propósito de dar respuesta a la diversidad de los perfiles de aprendizajes, se pueden aplicar de forma independiente. La primera consideración que debemos realizar es que previamente el docente debe conocer estos perfiles, sean mediante un cuestionario para recoger esta información de la propia voz de los alumnos, o mediante una observación de estos en diferentes contextos de aprendizaje (Morodo, 2017). Es crucial definir especialmente de qué manera aprende el grupo y cada alumno en concreto, para poder ajustar la evaluación.

La primera característica para conseguir una evaluación auténtica basada en los perfiles de aprendizaje detectados es variar el formato de comunicación y de respuesta de las tareas auténticas y ofrecer recursos para su resolución, siguiendo la metodología del diseño universal del aprendizaje que permite múltiples formas de representar, actuar y expresar y participar, como medida de atención a la diversidad (Sala, Sánchez, Giné y Díez, 2013). Esto conecta directamente con la característica basada en el realismo de las actividades, puesto que es muy habitual encontrar gran diversidad de estímulos y posibilidades para resolver una tarea, siendo el menos habitual explicar por escrito cómo se soluciona el problema. También es necesario ofrecer diversidad de recursos, que sirvan de ayuda a los alumnos y, a la vez, estén en sintonía con el contexto. En pocas ocasiones un profesional se encuentra delante de una tarea a resolver, y se le impide contrastar la información de la que dispone, consultar otras fuentes en caso de duda o compartir su proceso con otros profesionales. Así pues, programar diversos modos de recibir y comunicar el aprendizaje y poder acceder a diferentes recursos, fomenta unas condiciones de resolución realistas (Gulikers et al., 2004) y, a la vez, permite la participación de todos los alumnos.

Del hecho de ofrecer opciones y recursos diversos, nace la necesidad de ajustar la evaluación a la diversidad de capacidades de los alumnos sin vetar su participación en determinadas tareas que pueden resultar demasiado difíciles para que las resuelvan individualmente. Las prácticas de evaluación también deben basarse en

la diversificación y flexibilidad propios de una enseñanza adaptativa (Coll, Barberà y Onrubia, 2000). Para ello es preciso ofrecer ayudas (guías, actividades colaborativas, acceso a contenidos...), para que todos los alumnos puedan dar una respuesta con la que demuestren sus habilidades y competencias, sin rebajar la complejidad del reto. En este sentido, un acompañamiento durante el uso de estos recursos, en el que el docente ayude a conectar aprendizajes para superar lagunas, también resulta una acción importante para la evaluación inclusiva (Coll y Onrubia, 2002).

Lo cual, nos conduce a la última característica, que consiste en partir de las motivaciones de los alumnos. Está comprobado que las variables motivacionales condicionan el rendimiento académico de los estudiantes (Barca, Almeida, Porto, Peralbo y Brenlla, 2012). Por otro lado, la emoción facilita el aprendizaje; si los alumnos no conectan con la situación resulta mucho más difícil que se impliquen y que el aprendizaje sea significativo. A la vez, esta característica nos permite ajustar las propuestas contextualizadas y los aprendizajes que provocaremos a las realidades de los alumnos, lo cual repercutirá positivamente en su participación y resultados.

En conclusión, una evaluación auténtica debe ser inclusiva porque facilita a todos los aprendices diferentes propuestas, recursos, ayudas y situaciones para demostrar el desarrollo de sus competencias sin disminuir la complejidad y posibilitando de igual manera su aprendizaje. Todo ello, debe transmitir una práctica de evaluación segura, en la que los aprendices entiendan que cada uno tiene unas habilidades destacables, así como unas necesidades específicas, y que esto es lo que enriquece al grupo, proporcionando un ambiente social y emocionalmente seguro (Kaur, Noman y Awang-Hashim, 2015; Kaur, Noman y Nordin, 2017).

1.2.2. Experiencias de evaluación auténtica.

Como hemos comentado, la evaluación auténtica está constituida por una gran variedad de características que la conectan con la realidad, la orientan al aprendizaje y la vuelven sensible a la diversidad. Esto la convierte en una evaluación muy completa, pero compleja en el momento de implementar en el día

a día de las escolares y universitarias. Integrar sus distintas premisas en un único sistema de evaluación es una tarea difícil, especialmente para los docentes acostumbrados a prácticas de evaluación tradicionales (Villarroel et al., 2017). Incluso para los profesores más innovadores también supone un reto respetar todas sus características.

Sin embargo, existe un gran número de experiencias educativas en las que se utiliza la evaluación auténtica, tanto en ámbitos de educación obligatoria como no obligatoria y, especialmente, en relación con múltiples disciplinas (Frey et al., 2012). De forma general, en estas experiencias se aplican algunas de las características que la literatura señala como nucleares de la evaluación auténtica, como son el realismo, proponer tareas de alta complejidad cognitiva y ofrecer momentos de reflexión y de evaluación entre iguales (Fuentes y López, 2017; Koh, Tan y Tee, 2012; Martínez y Cantero, 2014; Meyers y Nutly, 2009; Raymond, Homer, Smith y Gray, 2013). Con el objetivo de mostrar algunas experiencias que nos puedan ayudar a diseñar, posteriormente, un programa de formación docente sobre evaluación auténtica, hemos seleccionado aquellas que suponen un ejemplo de la diversidad existente y de implementación en contextos específicos.

En primer lugar, analizaremos un grupo de experiencias sobre evaluación auténtica basadas en la transformación de las prácticas evaluativas más frecuentes de los docentes: las pruebas escritas. En segundo lugar, examinaremos un grupo de experiencias en el que se implementan prácticas propias de la evaluación auténtica en el sistema evaluativo de una asignatura.

Con relación a la autentificación de pruebas escritas, una propuesta interesante es la realizada por Villarroel, Boud, Bloxham, D. Bruna y C. Bruna (2019), en la que exponen un ciclo que permite transformar los exámenes escritos en prácticas de evaluación auténtica durante el propio proceso de evaluación. Coincidiendo con los momentos del proceso de evaluación, antes, durante y después, las acciones que se proponen son las siguientes: antes de la evaluación, una planificación para seleccionar lo que realmente importa, inyectar realismo en las pruebas y aumentar su complejidad; durante la evaluación, una aplicación de la prueba auténtica regulando sus condiciones como usar material, realizar respuestas colaborativas y

incluir simulaciones en entornos realistas profesionales; y, después de la evaluación, un despliegue de criterios de calificación consistente en involucrar a los estudiantes en tareas de coevaluación y usar la autoevaluación para juzgar los propios resultados.

Las investigadoras principales de esta propuesta la implementaron durante una formación a diecinueve profesores universitarios y analizaron los cambios que se producían (Villarroel, D. Bruna, Bustos, C. Bruna y Márquez, 2018). Los resultados del análisis de las pruebas antes y después de la formación revelaron que las pruebas escritas mejoraban especialmente en las dimensiones de autenticidad y complejidad cognitiva.

Otro conjunto de experiencias son las que se relacionan con el modelo de evaluación auténtica enunciado por Monereo (2009), en el que se propone a los docentes que evalúen sus evaluaciones para tomar consciencia de sus puntos débiles, de forma que puedan introducir cambios que las transformen progresivamente en pruebas más auténticas, que cumplen las siguientes características: a) que son realistas, es decir, que mantienen un grado de fidelidad con la realidad tanto el problema como su complejidad; b) que son relevantes, en tanto que resultan de utilidad para los estudiantes; c) que mantienen proximidad ecológica, es decir, que las propuestas se aproximan a las tareas y actividades del modelo de enseñanza habitual del docente y sus posibilidades reales de implementación; y d) que requieren una implicación identitaria, en tanto que permiten una socialización en contextos profesionales de forma progresiva. Este modelo se ha aplicado tanto en la educación superior (Mejías y Monereo, 2016; Monereo, Sánchez-Busqués y Suñé, 2012; Álvarez-Bernárdez y Monereo, 2016) como en la educación obligatoria (Monereo y Lemus, 2009; Monereo y Morcillo, 2011), hasta constituirse como sistema eficaz formativo de autenticación de las pruebas de evaluación tradicionales. En los próximos apartados profundizaremos en este modelo, denominado GAPPISA.

En segundo lugar, con relación al uso de prácticas de evaluación eficaces para una evaluación auténtica, las más efectivas son aquellas que permiten demostrar la resolución del problema y comunicarla, que están enfocadas a reunir diversas

evidencias de aprendizaje y que permitan al alumno puede participar en el análisis de sus progresos. Estamos hablando del portafolios, el análisis de casos, el planteamiento de problemas y las exposiciones de desempeño. Por sí solas estas actividades no hacen que una evaluación sea auténtica, pero sí facilitan el diseño de un sistema evaluativo orientado a las características de la evaluación auténtica más que las pruebas tradicionales. Además, combinadas en una misma programación, permiten a los estudiantes demostrar su aprendizaje de diferentes maneras (Darling-Hammond y Snyder, 2000). El uso de bases de orientaciones, rúbricas, guías de metacognición y listado de indicadores, entre otros instrumentos dirigidos a compartir, visibilizar y dirigir el aprendizaje, suelen asociarse a estas prácticas, facilitando claramente el desarrollo de muchas de las características de la evaluación auténtica.

De entre todas las prácticas de evaluación auténtica, quizá la que más resonancia ha tenido es el porfolio o carpeta de aprendizaje, que ha demostrado ser un instrumento altamente eficaz, especialmente en la educación superior (Barberà, Gewerc y Rodríguez, 2016; Díaz y Barroso, 2014; Guash et al., 2009). Al respecto, es interesante destacar el estudio de Sabirón y Arraiz (2013), que nos muestra la eficacia del uso del portafolio en el desarrollo y evaluación de competencias profesionales de estudiantes de Psicopedagogía y Magisterio, en una investigación realizada durante diez años con más de 600 estudiantes. De los resultados de la experiencia destacan el aumento de la implicación y compromiso de los estudiantes, y el hecho de que comprendan la evaluación como un instrumento para aprender.

De hecho, los beneficios de las prácticas de evaluación auténtica, más allá de su potencialidad para evaluar competencias y servir como instrumento de aprendizaje, también se reflejan en las opiniones de los alumnos. La evaluación tradicional no responde a las necesidades de los nuevos estudiantes y no permiten un aprendizaje profundo en tanto que no requieren de su participación (Kearney, 2013; Morales, 2010). Estudios recientes sobre los beneficios de una evaluación auténtica, apuntan que los estudiantes muestran satisfacción con estas prácticas, que definen como más accesibles, claras y útiles, a la vez que destacan que les

permiten mejorar sus calificaciones y la calidad de sus trabajos (Alderete y Edith, 2018).

1.2.3. La evaluación auténtica en el aula.

Después de conocer las experiencias de evaluación auténtica descritas, podemos afirmar que muchas escuelas y docentes introducen mejoras que les conducen a una evaluación auténtica de los aprendizajes según sus posibilidades de desarrollo institucional o profesional. Esto quiere decir que gran parte de los docentes introducen cambios relativos a algunas de las dimensiones como, por ejemplo, plantear situaciones realistas y/o compartir los criterios de calidad. El siguiente paso es diseñar un sistema de evaluación completo y homologable que permita su generalización y, a la vez, su contextualización. Estamos hablando de determinar todos los componentes y fases del proceso que tratamos en el apartado de evaluación de y en los procesos de enseñanza y aprendizaje.

Par ello, debemos determinar qué características serán las que definan las finalidades, el objeto, los referentes de valoración, los instrumentos, los momentos, los agentes y el proceso de evaluación para que sea un sistema de evaluación auténtica. Ahumada (2005) presenta una definición de estos componentes realizando el paralelo entre la evaluación tradicional y la evaluación auténtica (ver Tabla 12).

Tabla 12

Comparación entre el sistema de evaluación tradicional y el de evaluación auténtica

Componentes	Evaluación tradicional	Evaluación auténtica
Función principal	Certificar o calificar los aprendizajes.	Mejorar y orientar a los estudiantes en el proceso enseñanza-aprendizaje.
Relación con el aprendizaje	Paralela al proceso del aprendizaje.	Inherente o consustancial al aprender.
Información requerida	Evidencias concretas de logro de un aprendizaje.	Evidencias y vivencias personales.
Momento en que se realiza	Al finalizar un tema o una unidad.	Asociada a las actividades diarias de enseñanza y de aprendizaje.
Responsable principal	Procedimiento unidireccional externo al alumno.	Procedimiento colaborativo y multidireccional.

Análisis de los errores	Sancionan el error.	Reconocen el error y estimulan su superación.
Posibilidades de logro	Permite evaluar la adquisición de determinados conocimientos.	Permite evaluar competencias y desempeños.
Aprendizaje situado	Por lo general no le preocupa o desconoce en que ocurre el aprendizaje.	Considera los contextos en donde ocurren los aprendizajes.
Equidad en el trato	Distribuye a los alumnos en estratos creando jerarquías de excelencia.	Procura que todos los estudiantes aprendan a partir de su diversidad.
Reconocimiento al docente	Fuente principal de conocimiento.	Mediador entre los conocimientos previos y los nuevos.

Nota. Extraído de Ahumada (2005, p. 14)

Como podemos apreciar, en la tabla se detallan los componentes que hemos examinado en apartados anteriores, aunque algunos el autor los distribuye de diferente forma. Además, añade aspectos que hemos destacado como esenciales de la evaluación auténtica cuando describíamos sus características esenciales, como los que hacen referencia a su carácter inclusivo, por ejemplo, la equidad en el trato y el análisis de los errores; los respectivos a un enfoque competencial, por ejemplo, el aprendizaje situado y las posibilidades de logro; y los referentes a una evaluación formativa, por ejemplo, la relación con el aprendizaje y el momento en que se realiza. En definitiva, recoge las características de la evaluación auténtica señalando cómo dan respuesta a los componentes para diseñar un sistema de evaluación auténtica.

Sin embargo, al ser muchos los aspectos que el docente debe modificar en su práctica diaria para llegar a implementar una evaluación auténtica, entendemos que esta transformación debe realizarse progresivamente. Dada la interrelación entre los actuales enfoques sobre enseñanza y su conexión con la nueva manera de entender la evaluación, la introducción de cambios en alguno de los componentes provocará la necesidad de transformar otros. Por ejemplo, plantearse que el proceso de evaluación debe estar integrado en el de enseñanza, nos lleva directamente a cuestionarnos de qué manera recogeremos la información para evaluarla y qué información recogemos.

Por otro lado, es importante introducir unas mejoras u otras según las habilidades de los docentes y los recursos del centro. A la vez, debemos tener presente que transitar de una evaluación tradicional a una auténtica no consiste únicamente en aplicar algunas prácticas específicas, sino que es un proceso que supone transformar la cultura de evaluación en todos sus elementos. Si pretendemos que las innovaciones sean sostenibles en el tiempo, necesitamos un cambio sistémico que abarque toda la complejidad que representa la concreción de una nueva conceptualización de la evaluación. Se ha demostrado que construir comunidades de aprendizaje profesional en el centro favorece la sostenibilidad de las innovaciones educativas (Liesa et al., 2018). Por ello, es necesario repensar la evaluación conjuntamente y redefinir todos sus elementos para reposicionarse en una actitud evaluativa que permita construir colaborativamente un sistema de evaluación auténtica.

También debemos tener presentes las barreras existentes que dificultan esa transición. En primer lugar, representa un coste de recursos inicial para el diseño de nuevos materiales que muchos docentes no están dispuestos a asumir, generando esto una de las principales resistencias al cambio (Brown, 2015). En segundo lugar, el diseño curricular fragmentado en asignaturas y orientado, todavía en la práctica real de la mayoría de las escuelas, a los contenidos resulta una barrera para la implementación de una evaluación centrada en problemas reales que requieren, por definición, un trabajo interdisciplinar (Ahumada, 2005). En último lugar, quizá una de las resistencias más importante es, precisamente, las consecuencias que un cambio en la evaluación puede tener en las metodologías de aula y el enfoque de enseñanza en su totalidad. La evaluación suele producir efectos retroactivos en la enseñanza, de manera que los profesores que transforman su evaluación se ven impelidos a cuestionarse su metodología habitual cuando esta no está alineada con las dimensiones de la evaluación auténtica (Monereo, Castelló, Duran y Gómez, 2009). O, también, los efectos que tiene la evaluación sobre el aprendizaje, condicionando por ejemplo la forma de estudiar y de implicarse que tienen los estudiantes (Morales, 2009).

Efectivamente, poner a prueba el desempeño de los alumnos en problemas realistas y relevantes, ofrecer ayudas durante la evaluación para fomentar el

aprendizaje y saber que una evaluación justa no es la que es igual para todos sino aquella que se ajusta a las necesidades de cada alumno, obliga a los docentes a replantearse la enseñanza. La evaluación es el motor del aprendizaje (Sanmartí, 2007) y evaluar y enseñar son la cara de una misma moneda y, por lo tanto, deben estar en sintonía (Monereo, 2009). La evaluación auténtica también requiere una enseñanza competencial e inclusiva, apoyada en metodologías activas, para hacer real la máxima de que el alumno es el protagonista de su aprendizaje y que la evaluación solo es una estrategia más para que se cumpla ese mantra. Para que esto pase, es necesario que los docentes sean competentes en evaluación, evaluando sus evaluaciones para aprender a evaluar. En los siguientes apartados hablaremos de esta competencia evaluativa y de cómo desarrollarla a través de la formación inicial y continúa de los profesionales de la educación.

1.3. La Competencia en Evaluación de los Docentes

Hablar de competencias en educación no es nada nuevo como ya hemos señalado y, precisamente, fue en el ámbito profesional donde nació el término. Las competencias profesionales empezaron a tomar especial relevancia cuando en la selección de personal y gestión de equipos, en lugar de valorar únicamente los conocimientos que tenía la persona acreditados por una titulación, se empezó a dar importancia a cómo los ponían en práctica. Especialmente se buscaba que los trabajadores demostraran habilidades transversales a su trabajo como capacidad de trabajar en equipo, autonomía, sentido de la iniciativa y espíritu emprendedor, etc. Después se trasladaron estos requerimientos a la educación superior, transformándose en un contenido de aprendizaje y un enfoque pedagógico.

Dentro de este cambio, también surgió lo que se ha denominado las competencias clave que debe dominar el docente del siglo XXI (Perrenoud, 2004). Éstas condicionan el rol del profesor puesto que son el marco de actuación compartido y consensuado por los profesionales de la educación desde el que se puede definir en qué consiste la educación, de qué manera se lleva a cabo, qué implicaciones tiene, etc. Según Galvis (2007), uno de los primeros retos fue pasar de un perfil docente tradicional a uno basado en competencias para conseguir ser un profesional competente y, especialmente, para actuar como tal en todos los ámbitos en los que se desarrollan estas competencias: en el conocer, en el ser, en el convivir y en el hacer. Este desafío ha promovido el surgimiento de formaciones específicas y de modelos de gestión de la docencia basado en competencias dirigidos tanto a mejorar las competencias de los docentes como a saber enseñar por competencias (Pérez-Cabaní, Juandó y Argelagós, 2015)

Evidentemente, hay muchas y variadas propuestas en las que se despliegan las competencias en programas formativos para su desarrollo a lo largo de la etapa profesional, des de la formación inicial hasta la formación permanente. En todos ellos, sin embargo, la competencia en evaluación ha recibido poca atención (Sanahuja y Sánchez-Tarazaga, 2018). A continuación, explicaremos qué sucede en los programas de formación inicial de los futuros maestros respecto a la evaluación, para pasar a definir qué entendemos por competencia en evaluación y

cómo un cambio en esta remueve otros aspectos del rol docente, generando resistencias.

1.3.1. La importancia de una alfabetización en evaluación.

Como hemos apuntado anteriormente, una de las causas de la deficiencia en la competencia en evaluación que presentan los docentes es la formación inicial que estos reciben, en la que tanto la enseñanza sobre qué y cómo evaluar como las actividades con las que son evaluados suelen estar ancladas en una visión tradicional centrada más en la acreditación con la realización de exámenes como tarea principal. Estos son dos grandes retos para el desarrollo de la competencia docente en evaluación y es preciso analizar su origen.

La alfabetización en evaluación podemos definirla como el conjunto de saberes que tienen los docentes, después de vivir diversidad de experiencias al respecto, para implementar una evaluación y saber interpretar los resultados con la finalidad de mejorar (Stiggins, 2002; Popham, 2004). En la revisión de la literatura al respecto realizada por DeLuca, LaPointe-McEwan y Luhanga (2016) se define el término también equiparándolo con la competencia docente en evaluación. Sin embargo, en este estudio haremos una diferenciación entre los dos términos de la siguiente manera: la alfabetización en evaluación, como el conjunto de saberes y habilidades que son necesarias para implementar un sistema de evaluación, y ser competente en evaluación como las habilidades necesarias para poner en acción esos saberes de forma estratégica y situada; este último concepto lo desarrollaremos más en profundidad en el siguiente apartado.

Volviendo al término de alfabetización en evaluación de los docentes, la revisión teórica de DeLuca et al. (2016) sobre cuáles son los estándares y medidas internacionales utilizados, arroja luz sobre cuáles son las temáticas más abordadas dentro de este concepto. En su revisión de los documentos estatales sobre estándares de evaluación -normas gubernamentales y acreditación y certificaciones- de cinco países de habla inglesa y de la Europa continental, describen las siguientes temáticas nucleares: los propósitos de la evaluación, los procesos de evaluación, la comunicación de los resultados de la evaluación, la

equidad de la evaluación, la ética de la evaluación, la teoría de la medida, la evaluación para aprender y la evaluación educativa y el apoyo a los docentes.

El análisis de cómo están representadas estas temáticas a lo largo de los años muestra que ha habido una evolución en la importancia que se le da a cada una de las categorías desde el 1990 hasta la actualidad, a favor de un énfasis creciente en la evaluación para el aprendizaje. Y, respecto a los instrumentos analizados para medir la alfabetización en evaluación, cabe destacar que, en relación con las categorías señaladas, la menos representada es la de teoría de la medición. Estos resultados están en sintonía con el marco teórico presentado en el primer apartado, en el que destacábamos la importancia de una evaluación orientada al aprendizaje y describíamos los componentes y características del proceso evaluativo que los docentes deben conocer para llevarla a cabo.

Aunque parezca evidente, la importancia de que los profesores estén alfabetizados en evaluación es un hecho crucial para la educación y un requisito profesional, en tanto que los saberes sobre evaluación influyen en su práctica, de forma que un desconocimiento de la evaluación en el aula puede mermar la calidad de la enseñanza del docente (DeLuca, 2012). De hecho, se ha comprobado que una alfabetización de la evaluación tiene un impacto positivo en los resultados de los alumnos, especialmente en contraposición a aquellos docentes que no han estado alfabetizados (Mellati y Khademi, 2018). Sin embargo, la realidad de las aulas es que los docentes no se sienten capacitados, declaran que les faltan habilidades y encuentran dificultades para transformar sus prácticas de evaluación (Sanahuja y Sánchez-Tarazaga, 2018).

Por tanto, defendemos que la competencia en evaluación debe incluirse como competencia clave a desarrollar en los programas de formación inicial y continua y, además, evaluarse a través de instrumentos que permitan medir su desarrollo y el nivel de alfabetización. Es decir, debemos darle la importancia que merece en los planes formativos y en el desarrollo profesional del docente y no relegarla a un aprendizaje secundario o transversal que no se regula ni evalúa.

Por consiguiente, los estudios de formación docente deben contar obligatoriamente con temarios sobre por qué y de qué manera implementar la evaluación en la

enseñanza. Normalmente, sin embargo, después de una lectura rápida de las asignaturas obligatorias de las carreras de formación del profesorado, vemos que las asignaturas específicas sobre evaluación son escasas. La propuesta de trabajar la evaluación contextualizada en cada una de las didácticas específicas es correcta, sin embargo, dentro del complejo sistema que es la enseñanza universitaria, sin una programación interdepartamental muy cuidada que vele por el tratamiento de este tema, la evaluación como contenido tiende a diluirse dejando paso a los otros conocimientos centrales de la asignatura. Para que la evaluación no sea un aprendizaje residual y los docentes estén alfabetizados en esta, debe planificarse espacios y tiempos propios para la evaluación en los planes de estudio y evaluar su competencia como tal (Martínez, 2013).

En consecuencia, la forma en la que se evalúa a los futuros docentes debe ser también formativa y centrada en el aprendizaje, tanto durante los planes de estudio como en las pruebas de acceso a la profesión docente. Respecto a cómo se evalúa a los futuros maestros, la "cultura del test" continúa siendo casi un imperativo en las prácticas de evaluación en la formación inicial del profesorado, dirigida a medir y acreditar, la perspectiva de los futuros maestros sobre qué significa la evaluación así lo demuestra. Existen estudios que corroboran que los alumnos se implican más en su aprendizaje cuando participan en dinámicas de evaluación en las que tienen voz y que, por lo tanto, la evaluación sumativa no es suficiente (Gómez-Devís, García-Raffi y Saneluterio, 2017). También se han diseñado cuestionarios para evaluar la percepción de los estudiantes y analizar si tiende hacia un enfoque superficial o uno profundo (Doménech y Fortea, 2015). Instrumentos como este son una buena estrategia para promover la reflexión sobre de los docentes universitarios.

Lo importante, es que se deben superar las contradicciones entre lo que se predica en el temario de formación inicial de maestros -necesidad implementar una evaluación formativa- y lo que se aplica -una evaluación mayoritariamente acreditativa- en tanto que es necesario predicar con el ejemplo. Como señalan López-Pastor y Pérez-Pueyo (2017):

Existen algunos trabajos (Fernández-Pérez, 1989; López- Pastor, 2009, 2013; Lorente & Kirk, 2013, Palacios & López-Pastor, 2013; Teje- dor, 1998; Tonucci, 2010) que denuncian las graves consecuencias que tiene la habitual incoherencia que se repite año tras año en FIP, y que consiste en recibir discursos sobre cómo hacer la evaluación en la práctica educativa que son completamente contrarios con la evaluación que se lleva a cabo en dichas asignaturas. (p. 52)

Así pues, es necesario también formar a los docentes universitarios, los cuales, además, están también en proceso de adaptarse a los retos de la Declaración de Bolonia, que demanda también una evaluación de competencias profesionales. En el contexto catalán, por ejemplo, según el estudio de Ion y Cano (2012), todavía no hay propuestas formativas suficientes que ayuden a los docentes universitarios a implementar una evaluación por competencias, aunque el interés es cada vez mayor.

Por otro lado, Clark y Rust (2006) también descubrieron que los exámenes seguían siendo el instrumento principal cuando quisieron establecer un enfoque de evaluación centrado en el aprendizaje y encontraron resistencias notables por parte de todos los agentes. Pero lo interesante de su estudio es que presentaron un modelo que consistía en analizar las propias actividades de evaluación desde cinco puntos de vista -futuros maestros, facultad, administraciones, maestros colaboradoras y alumnos de escuela-, para describir sus consecuencias en los tres momentos del proceso. El modelo tenía el propósito de realizar cambios para que los docentes involucraran a los futuros maestros en la evaluación. La colaboración entre universidad y centros educativos es también una buena estrategia formativa.

Respecto a las pruebas de acceso a la profesión, existen estudios (Goodman, Arbona y Domínguez, 2008) que analizan la relación entre las pruebas clásicas, los exámenes finales que imponen las instituciones y las actividades de evaluación relacionadas con el rendimiento como los portfolios o cuestionarios de atributos profesionales. Los resultados ponen en duda la efectividad de los exámenes estandarizados como medida de selección de maestros realmente competentes.

En conclusión, debemos alfabetizar en evaluación a los futuros docentes a través de la formación sobre evaluación en los planes de estudio como competencia a desarrollar y, para ello, debemos realizar ya en la formación inicial prácticas de evaluación formativa centradas en ayudar a los futuros maestros a aprender y desarrollar sus competencias. Existen experiencias positivas al respecto, tanto en las facultades de formación del profesorado como en la enseñanza universitaria en general (Cano y Ion, 2016; Blanch et al., 2011).

1.3.2. La necesidad de ser competentes en evaluación.

Una alfabetización en evaluación es un primer paso, sin embargo, no garantiza que luego los docentes sean efectivos en este ámbito o que enfoquen sus habilidades hacia una evaluación orientada al aprendizaje. De hecho, mejorar el conocimiento sobre evaluación es necesario, pero, en la puesta en práctica, las concepciones pueden seguir siendo opuestas a la alfabetización recibida (Deneen y Brown, 2016) que se sostienen sobre evaluación pueden no haber mejorado. Para ser competente en evaluación es necesario practicar en los contextos reales de enseñanza.

Además, ser competente en evaluación no solo consiste en evaluar a los alumnos, sino según Tonda y Medina (2013), significa también evaluar la función docente, los programas y recursos y las propias evaluaciones, por lo tanto, implica poner en acción todos los saberes adquiridos en la alfabetización y valorar su funcionalidad en un contexto específico. Al respecto, Sanahuja y Sánchez-Tarazaga (2018) también distinguen diferentes niveles de consecución de esta competencia que muestran la importancia del desempeño como medida de la competencia; así destaca que es necesario dominar el conocimiento, que significa saber definir, la explicación, que se refiere a saber cómo hacerlo, la simulación, como el hecho de demostrar cómo se hace y el desempeño real, que significa hacerlo.

A partir de esas conceptualizaciones, definimos la competencia en evaluación del docente como un acto estratégico en el que el profesional es capaz de promover y resolver diversidad de situaciones evaluativas, es decir, aquél que es capaz de planificar su evaluación según la finalidad y la situación, implementarla haciendo

los ajustes necesarios *in situ* y evaluar su puesta en práctica para poder mejorar. Según Monereo y Castelló (2009) esta competencia está formada por un conjunto de subcompetencias que definen lo que significa que un docente sepa, sepa hacer y sepa ser en relación a la evaluación; estas son: ser capaz de delimitar el contexto de la evaluación, ser capaz de seleccionar y construir la metodología de evaluación, ser capaz de regular la propia actuación, ser capaz de interpretar los resultados y realizar una metaevaluación y, por último, ser capaz de comunicar los resultados.

Dichas subcompetencias se concretan en una serie de actividades que el docente debe llevar a cabo y que pueden guiar la construcción, por ejemplo, de una programación formativa al respecto. Estas subcompetencias ya se han utilizado en un estudio en el que, Y el estudio de Tonda y Medina (2013) examinan el nivel de competencia evaluativa en una muestra de 100 docentes de educación secundaria. Los resultados nos indican que las que muestran menor aceptación en los profesores en activo son las relativas a la selección de instrumentos y construcción de metodologías y la delimitación del contexto de la evaluación.

A propósito, ser competente en evaluación consiste en ser estratégico en el uso de las siguientes habilidades: uno, ser capaz de planificar la evaluación según el contexto y la situación, dos, aplicar el proceso de evaluación de forma estratégica poniendo en práctica los ajustes necesarios y, tres, ser capaz de autoevaluar la implementación y extraer información para mejorar futuras intervenciones evaluativas.

En primer lugar, cuando hablamos de ser capaz de planificar la evaluación según el contexto y la situación, queremos decir, por un lado, saber diseñar un sistema de evaluación siguiendo los componentes y el proceso que detallábamos en el primer apartado de este capítulo y, por otro lado, ajustarlo al contexto considerando las posibilidades del docente y la situación educativa.

Recordemos que, programar el proceso de evaluación significa realizar un conjunto de tareas que van desde la definición de criterios hasta el análisis del propio proceso de evaluación. No se trata únicamente de diseñar una prueba centrada en los contenidos de la disciplina y en corregirla para calcular en qué medida cada alumno ha adquirido los conocimientos, como tradicionalmente se ha hecho. A su

vez, debe hacerse desde la visión que compartimos desde el inicio del estudio, enfocándose hacia una evaluación orientada al aprendizaje.

En segundo lugar, cuando hablamos de aplicar el proceso de evaluación de forma estratégica poniendo en práctica los ajustes necesarios, queremos decir activar los dispositivos necesarios para que la evaluación no se aleje de nuestra planificación y, a la vez, dé respuesta a la realidad del aula, especialmente a las necesidades de nuestros alumnos. Esto incluye, si es necesario, modificar algunos de los planteamientos iniciales, por muy profundos que nos parezcan. Mateo (2000) por ejemplo, señala que el docente debe ser capaz de valorar si se puede aplicar según el momento y las necesidades detectadas. Al fin y al cabo, son los procesos de enseñanza y aprendizaje una compleja y orgánica actividad.

En tercer lugar, cuando hablamos ser capaz de autoevaluar la evaluación y extraer información para mejorar la enseñanza, debemos interpretarlo como el momento en el que los docentes aprenden de su propia actuación. Primero, con el objetivo de valorar si el proceso de evaluación es correcto y, segundo, para evaluar qué mejoras debería introducir en su enseñanza para obtener resultados mejores.

En este sentido, es interesante destacar el listado de los siete criterios formulados por López-Pastor y Pérez-Pueyo (2017) para evaluar un sistema de evaluación y valorar si es formativo y compartido. Dichos criterios, que funcionan en calidad de principios que orientarla toma de decisiones, son los siguientes: uno, el criterio de adecuación respecto al diseño curricular, a las características del alumnado y el contexto y respecto a los planteamientos educativos; dos, el criterio de relevancia según el cual hay que seleccionar aquello a evaluar; tres, el criterio de veracidad que vela por el rigor, la credibilidad y fiabilidad; cuatro, el criterio formativo que da cuenta de la función de la evaluación para mejorar los procesos de enseñanza y evaluación; cinco, el criterio de integración de la evaluación en los procesos cotidianos de enseñanza, de los alumnos en los procesos de evaluación y de contenidos y competencias; seis, el criterio de viabilidad según el cual el sistema de evaluación debe ser posible aplicarlo considerando las condiciones reales de cada docente; y siete, el criterio ético tanto en el tratamiento de la información como en el uso de poder que implica la calificación. Estos criterios pueden considerarse

también en el momento de planificar la evaluación, incluso en el durante como medio de evaluación.

Sin embargo, los estudios nos muestran que los docentes no suelen ser activos respecto a la metaevaluación, más allá de realizar una autorreflexión rápida sin criterios establecidos previamente (Tonda y Medina, 2013). Esta omisión tiene dos consecuencias, a nuestro parecer, ambas negativas y perjudiciales para el sistema de evaluación en sí. Por un lado, la evaluación pierde una de sus funcionalidades más importante: la mejora de la práctica educativa. El estudio de Smith, Rodgers, Hawk y Fulcher (2015) demuestra que los docentes, después de la evaluación, raramente interpretan los resultados con el fin de introducir nuevas acciones educativas dirigidas a resolver las dificultades detectadas, y mucho menos con el objetivo para volver a evaluar a los alumnos permitiéndoles obtener nuevos resultados. Esto tiene efectos directos en la eficacia del sistema de evaluación formativo. Por ejemplo, provoca que los alumnos perciban que la evaluación no tiene mucha importancia y, por tanto, no sea un medio para aprender. Esta falta de metaevaluación sucede desde la formación inicial de los docentes, lo que tiene como consecuencia que los futuros maestros ya carezcan de la competencia en autoevaluarse (Gallego y Rodríguez, 2017).

En definitiva, ser competente en evaluación quiere decir, por un lado, tener claro cómo planificar, implementar y evaluar un sistema de evaluación y, por otro lado, implementarlo y superar los incidentes que surjan, experimentando este proceso como una oportunidad de aprendizaje. Todo ello, en una realidad en la que los docentes se encontrarán con el reto añadido de tener que enfrentarse a la desconexión entre la alfabetización en evaluación, desde la formación inicial, y la práctica real en el aula (Deneen y Brown, 2016). Para superar las dificultades, deberán pasar por un proceso de reaprendizaje de lo que significa la evaluación, y tener presente que la complejidad y el dinamismo de una evaluación orientada al aprendizaje hará emerger desafíos de forma continuada. Llegados a este punto, es casi imperativo destacar la importancia que la formación continua tendrá en el ámbito de la evaluación, presentándose como una alternativa necesaria para suplir las carencias de la alfabetización en evaluación y formar docentes que puedan

enfrentar estos desafíos de forma competente. En el siguiente apartado desarrollaremos propuestas eficaces de intervención en esta línea.

1.4. La formación Continua sobre Evaluación

El aprendizaje a lo largo y ancho de la vida es un término que da cuenta de que el desarrollo de las personas implica aprendizaje constante y que no se puede equiparar únicamente a la formación reglada, sino que a lo largo de la vida las personas participamos en experiencias educativas (cursos, talleres, charlas, etc.) que nos permiten conocer y desarrollar nuevos conocimientos y habilidades. La formación continua es uno de los pilares de este aprendizaje a lo largo de la vida y los profesionales de la educación se sirven constantemente de esta herramienta para mejorar sus prácticas profesionales.

Es innegable que la profesión docente, encargada de formar a los futuros ciudadanos del planeta, tiene casi como imperante revisar constantemente sus prácticas, incluso a veces, sus motivos para ajustarse a la sociedad a la que se debe. Así pues, la formación permanente del profesorado cumple una función indispensable en este sentido y son muchas y muy variadas las propuestas formativas, tanto en lo que se refiere a la temática que tratan como a la manera de abordarla, consiguiendo en menor o mayor medida la transferencia de los aprendizajes a la práctica educativa diaria.

Dentro de este conjunto de acciones formativas, creemos que actualmente, las propuestas sobre evaluación son de especial relevancia en el sistema educativo actual que está cada vez más posicionado en un enfoque por competencias. Por un lado, como hemos expuesto en los primeros apartados de este capítulo, se está forjando un cambio de paradigma evaluativo que rompe con la forma tradicional de evaluar y tiene efectos retroactivos en la enseñanza y el aprendizaje. Por lo tanto, un cambio de considerable envergadura para maestros y docentes, porque, a la vez que se tiende hacia un nuevo enfoque, éste todavía está lleno de ambigüedad y falta de consenso en la comunidad educativa, de forma que en muchos centros educativos y, especialmente para muchos docentes, el cambio está siendo lento y dificultoso (Moreno, 2012), y para otros el enfoque de competencias no es un motivo suficiente para modificar sus prácticas (Tonda y Medina, 2013).

Por otro lado, tal y como hemos justificado en el apartado anterior, la deficiente alfabetización sobre evaluación de los docentes producida por una formación inicial deficiente, se traduce en el desarrollo de una pobre competencia evaluativa. Esto tiene como consecuencia que el colectivo docente normalmente no tenga las estrategias necesarias como para adaptar su evaluación al nuevo paradigma. Y, aún habiendo vivido sistemas de evaluación formativa durante su formación inicial, los estudiantes declaran necesitar apoyo para poder implementarla con éxito, normalmente, para superar las dificultades y resistencias del centro (Hamodi y López, 2012). Bajo esta dificultad nace un gran número de propuestas formativas sobre evaluación: conferencias, seminarios, cursos, charlas, congresos, etc. con el objetivo de suplir estas carencias.

Por todo esto, se justifica con creces la necesidad de dedicar nuestros esfuerzos a la formación continua de los docentes en la tan importante área de la evaluación. Porque precisamente, como afirman Tonda y Medina (2013), la mejor manera de desarrollar la competencia evaluativa es participar en formaciones continuas basadas en la investigación colaborativa e innovación educativa. Sin embargo, no todas las estrategias formativas impactan de igual manera en el desarrollo profesional del docente. Por eso, en este apartado, primeramente, expondremos un breve análisis de cuáles son las características que, a nuestro juicio, definen un programa de formación docente de calidad. Seguidamente, describiremos algunas propuestas formativas sobre evaluación para, por último, presentar al detalle la formación que proponemos, experiencia de la cual parte nuestra investigación.

1.4.1. La formación permanente y su transferencia.

Las actividades de formación permanente en las que participan los docentes pueden materializarse en diferentes experiencias: cursos y seminarios, charlas, conferencias, jornadas o congresos, proyectos de innovación pedagógica del propio centro o institución educativa, reflexión colaborativa en grupos de trabajo intercentros o del propio centro, etc. y hay que saber seleccionar la modalidad en función del propósito que se persiga. Además, se tienen que ajustar las condiciones de esta también, los agentes que intervienen, la temporalización, el tipo de acciones

o actividades, los recursos, la programación de contenidos, etc. todo ello puede determinar la consecución o no de los objetivos formativos.

El estudio que presentaremos en el siguiente capítulo, como ya hemos señalado, parte del propósito de implementar una formación sobre evaluación auténtica que impacte en la realidad educativa de los docentes y perdure en el tiempo. Por ello, es importante para nosotros escoger modalidades formativas que promuevan la transferencia de los aprendizajes a las prácticas docentes.

Por ello, hemos identificado cómo debería ser una formación de calidad. El estudio de Comas (2019) titulado *Sirve la formación permanente del profesorado para mejorar los resultados educativos del alumnado?*, investiga cuáles son las características óptimas de los programas de formación permanente del profesorado para que estos tengan un impacto en los resultados de los alumnos. Aunque la autora señala que los resultados no son conclusivos, hay indicios suficientes como para realizar algunas recomendaciones sobre los aspectos esenciales. A continuación, realizamos un listado de estas características, apoyándonos también otros estudios con el fin de tenerlas presentes para el diseño de nuestro proyecto de formación sobre evaluación. Las características, basadas en el análisis de Comas (2019) son las siguientes:

- *Partir de la realidad del docente (sus prácticas, sus necesidades, sus creencias, etc.) vinculando la formación a ésta.* Es importante que la formación dé respuesta a las demandas de los docentes, pero más importante es que conecte con su realidad, de forma que se haga explícita la manera en la que éstos pueden aplicar los aprendizajes en su contexto actual: los recursos de los que disponen, la metodología que utilizan normalmente, las dinámicas de equipo, etc. No se trata de que la formación reproduzca o ensalce aquellos aspectos que podrían mejorar dándolos por buenos, se trata de situarse en lo que Vygotsky (1978) denominó la zona de desarrollo próximo, en este caso de los docentes, con el fin de partir de sus prácticas cotidianas para moverlo a una etapa posterior.
- *Permitir actividades en las que se pueda reflexionar sobre la propia práctica.* Estrechamente relacionado con el punto anterior, además de conectar con

el contexto real de los docentes es necesario que éstos puedan analizar, exponer y discutir cómo los conocimientos y habilidades que están aprendiendo se conectan con su práctica. Los docentes suelen tener menos tiempo del que necesitan para encontrarse entre ellos y reflexionar sobre su enseñanza junto a otros y con cierta distancia. Las formaciones no son solo un buen momento para ello, sino que, además, ofrecer la oportunidad de hacerlo facilita conectar los nuevos saberes con los propios, así como descubrir nuevas conexiones y generar ideas que permitan innovar en la propia práctica.

- *Programar un ciclo formativo completo con fases claras e interconectadas.* Esta característica hace referencia a la idea de no concebir la formación únicamente con una trasmisión de contenido, sino con un proceso de asesoramiento. Se trata de implementar una propuesta que vaya desde la reflexión hasta la evaluación de la puesta en práctica. Diversas experiencias presentadas en el estudio de López-Pastor y Pérez-Pueyo (2017) demuestran la eficacia del siguiente ciclo: a) elaborar unos criterios de calidad evaluativa, b) fijar unas líneas de actuación que faciliten encontrar el sistema de evaluación adecuado, c) ponerlo en práctica e iniciar ciclos de investigación-acción para ir mejorando progresivamente. Las experiencias prácticas que presentan en el estudio demuestran que este ciclo es una muy buena estrategia para sostener los resultados de la formación e ir mejorándolos.
- *Diseñar actividades de aprendizaje basadas en la acción.* Como en el caso de los alumnos, es importante que el aprendizaje sea significativo y para ello es recomendable implicar a los docentes llamándolos a la acción. Pueden participar en actividades como análisis de documentos, dinámicas de construcción de conocimiento, observación de clases, planificación de unidades didácticas, diseños de instrumentos, debates, feedback entre expertos, etc. La elección de las actividades dependerá del propósito y el contenido de cada formación, siendo recomendable elegir aquellas que sean realistas o que tengan como resultado productos que después puedan utilizar en su día a día.

- *Programar tiempo real para poner en práctica los aprendizajes y acompañar esta aplicación.* Esta es una característica indispensable si lo que se quiere es conseguir una transferencia real de los aprendizajes. Está directamente ligada con la idea de aprender mediante la acción. La formación debe planificar un tiempo exclusivamente para que el docente pueda probar lo que aprendió, liberándolo, si es preciso, de otras tareas, por lo menos hasta que éste domine el nuevo aprendizaje de forma mínimamente eficiente. Si no es así, el coste de la nueva implementación será muy elevado para el docente. Además, el formador, asesor o compañero más experimentado, no debería desaparecer una vez se empiezan a dominar los conocimientos, porque es precisamente en la aplicación de estos en su práctica educativa real donde los docentes necesitan más ayuda. Se debería acompañar al docente al menos en los primeros momentos, que son claves pues es cuando este acabará de ajustar el aprendizaje a su idiosincrasia. Las prácticas de *coaching* se han visto altamente efectivas en este sentido (Johnson, Finlon, Kobak y Izard, 2017; Kraft, Blazar y Hogan, 2018). De esta manera, además, aseguramos la calidad de las prácticas.
- *Disponer de un liderazgo pedagógico que vele por la construcción de una comunidad educativa que aprende y se co-responsabiliza.* Es tarea del equipo directivo junto con los líderes pedagógicos (cuando estos no forman parte del equipo directivo), promover acciones formativas que den respuesta a las necesidades detectadas en el centro. Además, se ha demostrado que la implicación del equipo directivo en las mismas formaciones influye en los resultados positivamente. Junto con esto, es necesario la creación de una comunidad educativa que aprenda, donde la responsabilidad de los resultados y del bienestar de los alumnos es compartida entre todos los docentes (Robinson, Hohepa y Lloyd, 2009). El líder pedagógico debe facilitar espacios y tiempos para que se constituyan y desarrollen comisiones o equipos de trabajo centrados en llevar a la práctica aquellas innovaciones que se han trabajado en las diferentes formaciones.
- *Promover y asegurar la transferencia de los aprendizajes en el tiempo, a través de la autoevaluación de la implementación y la evaluación del impacto.* El último paso del ciclo consiste en dotar a los docentes de

herramientas para que puedan continuar con la implementación. Y en este proceso, la evaluación de la formación y su impacto es un aspecto crucial, pues, como afirma Iranzo (2011): "no se puede acompañar hacia la mejora educativa sin construir conocimiento situado con motivo de la apropiación del proceso por parte de sus protagonistas y para esta apropiación" (p. 36). Por ello, se deben diseñar mecanismos que faciliten la autonomía de los docentes para autoevaluarse (establecimiento de grupos de trabajos e instrumentos como tablas de observación, rúbricas, etc.) y evaluar el impacto de sus innovaciones (colaboración con grupos de investigación, realización de jornadas, etc.). Este último punto es altamente motivante.

Pensamos que todas estas características se han de cumplir en mayor o menor medida para asegurar que la formación tiene un impacto en la realidad educativa de los alumnos. Pero, para este fin en concreto, la última nos parece de especial relevancia para que la formación permanente cumpla sus fines. Porque, aunque ciertamente ninguna formación pasa desapercibida, también hay algunos estudios que nos indican que todavía no hay resultados generales que aseguren que la formación permanente tiene realmente un impacto en la mejora de las escuelas (Iranzo, 2011; Imbernón, 2007). Estos mismos autores señalan que el motivo de que no conozcamos esta influencia se debe a que habitualmente no se realiza una evaluación de los efectos de la formación en las prácticas de los docentes.

En este sentido, pensamos que la colaboración entre grupos de investigación académica y centros educativos puede ser una estrategia eficaz para generar conocimiento que apoye la formación permanente y las innovaciones docentes.

1.4.2. La formación sobre evaluación.

En la última década ha resurgido con fuerza el tema de evaluación, siendo un tema recursivo en muchas formaciones y proyectos de innovación educativa, con su correspondiente incremento en investigación sobre el tema. Creemos que el hecho de que, entre todos los temas de formación sobre evaluación los docentes demanden formaciones sobre métodos, instrumentos, actividades y planificación de estas, puede responder a tres motivos: un primer motivo, que ya hemos

apuntado en el apartado anterior, se debe a que de entre todas las subcompetencias que deben dominar, las que hacen referencia a la selección de instrumentos y construcción de metodologías y delimitación del contexto de la evaluación, es la menos desarrollada, según el estudio de Tonda y Medina (2013) sobre competencia evaluativa.

Otro motivo, coincide con el hecho de que sean estas actividades las que les influyen de manera recursiva y a las que están obligados para poder recoger los resultados, como mínimo para medirlos, cuantificarlos y entregar el boletín de notas al que están obligados por ley. Por contra, como comentamos cuando hablábamos de las fases a seguir en la planificación del sistema de evaluación de competencias, el resto de las fases más globales y de aplicación no tan inmediata, se olvidan.

El tercer motivo, que también se deduce de lo estudiado hasta el momento, es que, con la implementación del enfoque competencial en la enseñanza, los docentes se ven obligados a aprender estrategias de evaluación que no han visto hasta el momento. Un número elevado de docentes en activo hace más de 10 e incluso 20 años que realizaron la formación inicial, así pues, deben actualizarse constantemente sobre muchos temas. Pero enseñanza y evaluación, dentro del nuevo enfoque, son las dos caras de una misma moneda y, por tanto, formarse en evaluación resulta un imperativo.

Las formaciones deben dar respuesta a estas necesidades declaradas por los docentes, pero creemos que siempre, además, deben ir un poco más allá, obligándoles a replantearse las finalidades de la evaluación, los procesos necesarios para que la innovación sea factible o las tareas necesarias para evaluar la calidad de la propia evaluación, entre otros. Especialmente para desarrollar la competencia evaluativa. En el estudio de Tonda y Medina (2013) al que nos hemos referido con anterioridad, se demuestra que la formación permanente que reciben los docentes influye en sus actitudes respecto la evaluación. Los autores relacionan las subcompetencias evaluativas de los docentes con el tipo de formaciones que realizan, y obtienen que las formaciones basadas en la indagación e innovación educativa y las basadas en el trabajo colaborativo realizado mediante grupos de

profesores del centro tienen actitudes más positivas hacia la mayoría de las subcompetencias.

Así pues, es preciso analizar algunas propuestas de formación continua sobre evaluación y de proyectos formativos sobre evaluación, con el fin de destacar sus bondades y aprender de sus limitaciones. A continuación, expondremos algunas investigaciones que demuestran los efectos que las formaciones sobre evaluación para el aprendizaje han tenido en docentes, alumnos y centros escolares, para destacar algunas de sus bondades.

1.4.2.1. Maestros que desarrollan evaluaciones para el aprendizaje: impacto en el logro estudiantil.

La experiencia formativa de William, Lee, Harrison y Black (2004) parte de la premisa de que la metodología basada en que el investigador o el asesor vaya a la escuela y diga lo que los docentes deben hacer, normalmente no se traduce en una puesta en práctica de la investigación. La complejidad de la propia acción supone un reto tan elevado para la investigación, que parece casi imposible que ésta pueda definir exactamente lo que hay que hacer. Así pues, la formación se basó en un trabajo colaborativo entre un grupo de 24 maestros de seis escuelas distintas y unos asesores que sugerían direcciones para que cada docente encontrara la forma de implementar los principios generales de la evaluación formativa en el aula.

La intervención se concretó en: un conjunto de sesiones en las que se presentaban los principios de la evaluación formativa y, en base a eso, los docentes diseñaban sus propios planes partiendo de lo que ya hacían y unas sesiones de seguimiento durante seis meses con el asesor mientras realizaban la implementación, para analizarla y planificar mejoras.

Para comprobar la eficacia de la formación se analizaron los resultados en evaluaciones externas de los estudiantes que habían participado en una evaluación formativa, en comparación con otros que no lo habían hecho, obteniendo los primeros mejores resultados de forma significativa. Así pues, la formación resultó ser efectiva; además, la mayoría de los maestros declaran querer sostener las prácticas de evaluación formativas.

1.4.2.2. Entendiendo cómo los profesores se involucran en la evaluación formativa.

En el artículo de Sondergeld et al. (2010) se presentan los resultados de un programa de desarrollo profesional sostenido centrado en la evaluación formativa, llamado *Keeping Learning on Track Program* (a partir de ahora KLT), de dos años de duración. Éste tenía como objetivo que los maestros recogieran evidencias de aprendizaje día a día y las analizaran para adaptar su enseñanza y resolver las dificultades observadas en la mayor brevedad posible, es decir, utilizar la evaluación para regular el aprendizaje.

Participaron como muestra del estudio 65 maestros de forma voluntaria, los cuáles realizaron las siguientes actividades propias del programa: taller inicial de dos días donde se presentaba el programa y sus lineamientos teóricos y hasta siete reuniones mensuales con los asesores y los participantes en las que se analizaba las implementaciones (compartían reflexiones, estrategias, analizaban evidencias, etc.).

Los resultados positivos son múltiples y todos los maestros relataron que la evaluación formativa modificó sus prácticas profesionales y el aprendizaje de los alumnos, a destacar los siguientes resultados: notas más altas en las pruebas estandarizadas, más información sobre contenido evaluado (por la recogida de evidencias), aumento en la participación, motivación y responsabilidad hacia su trabajo por parte de los estudiantes. Pero uno de los aspectos más destacable, es que la implementación produjo comentarios positivos de los alumnos, hecho que sostuvo la motivación e implicación de los docentes por continuar mejorando su evaluación. Otro aspecto es que las reuniones propuestas proporcionaron un ambiente seguro de aprendizaje, un grupo de apoyo basado en la práctica guiada favoreció la reflexión y anticiparse a ciertas dificultades.

1.4.2.3. Evaluación formativa y compartida en educación: experiencias de éxito en todas las etapas educativas.

También es destacable el libro coordinado por López-Pastor y Pérez-Pueyo (2017) en el que se muestran experiencias de éxito en la implementación de una evaluación formativa y compartida en todas las etapas educativas. Este libro es el resultado de 20 años de trabajo para el diseño y desarrollo de un sistema de evaluación formativa y compartida y el análisis su puesta en práctica en todos los niveles educativos, resultando el libro como una gran base de datos de experiencias reales que anima a otros docentes a tomar el ejemplo y transformar la evaluación.

La propuesta general, como ya expusimos en el punto anterior cuando hablábamos de las características, se basa en los siguientes pasos sucesivos entre ellos: primero, elaborar unos criterios de calidad evaluativa; segundo, fijar unas líneas de actuación que faciliten encontrar el sistema de evaluación adecuado; y, tercero, ponerlo en práctica e iniciar ciclos de investigación-acción para ir mejorando progresivamente.

1.4.2.4. Evaluar para aprender: la evaluación al servicio del aprendizaje del alumno.

Por último, también es preciso señalar que existen planes de formación permanente impulsados por los departamentos de educación de diferentes países, así como centro de recursos pedagógicos que ponen a disposición de los docentes formaciones y recursos sobre las necesidades que tengan. Entre estas propuestas, queremos destacar *la Xarxa de Competències Bàsiques*, una red de escuelas en formación y desarrollo profesional docente basada en el aprendizaje entre iguales, que tiene como objetivo transformar la escuela en una organización que aprende, específicamente, sobre el aprendizaje y la evaluación por competencias.

La primera temática tratada fue sobre evaluación, *Avaluar per aprendre: l'avaluació al servei de l'aprenentatge de l'alumnat*, y, en torno a este tema se realizaron acciones como formaciones internas de centro, jornadas, seminarios de trabajo, construcción de un banco de materiales, etc. Todas estas acciones bajo la

metodología de la práctica reflexiva mediante ciclos de acción-reflexión-acción. Las fases de este ciclo son: uno, diagnóstico y establecimiento de objetivos de mejora compartidos, criterios e indicadores de evaluación; dos, análisis crítica de la propia práctica; tres, enriquecimiento gradual de las propias prácticas o estrategia de aula; cuatro, materialización contextualizada de nuevas prácticas o estrategias de aula; cinco, experimentación de las nuevas prácticas en diferentes aulas y observación de los efectos en los aprendizajes de los alumnos; seis, valoración global de los resultados de la experimentación; y siete, evaluación de la mejora y análisis de la sostenibilidad y la transferencia de las acciones (Xarxa de Competències Bàsiques, 2019). Y, aunque todavía no existen estudios específicos sobre los resultados, en las jornadas anuales se presentan las experiencias reales a través de comunicaciones y las ponencias sobre los temas a tratar¹.

Todas las experiencias que hemos descrito tienen resultados positivos en los aprendizajes de los alumnos, lo cual demuestra que la evaluación para el aprendizaje es necesaria. Además, las formaciones cumplen, en mayor o menor medida, las características que hemos destacado en puntos anteriores sobre la formación permanente para conseguir una transferencia real. También, por la naturaleza de las propuestas formativas, basadas en la reflexión sobre la práctica, el andamiaje sostenido gracias a un grupo de apoyo (otros docentes y expertos en la temática), una temporalización de calidad y una evaluación de la implementación, pensamos que todas permiten el desarrollo de la competencia evaluativa.

Además, entrando en detalle, queremos destacar algunas bondades de las propuestas. De la primera formación resaltamos la figura del asesor como guía y no como dador de conocimientos (William et al., 2004). Respecto a la segunda formación destacamos la importancia de que los docentes sean conscientes de sus

¹ La primera jornada se titulaba, **IX Jornada anual Congrés europeu. Avaluat per aprendre** (Barcelona, 12 i 13 de maig de 2011); y las ponencias y comunicaciones se encuentran en:

<http://xtec.gencat.cat/ca/curriculum/xarxacb/jornades/jornada2011> La segunda jornada se titulaba, **X Jornada anual. L'avaluació per a la millora de l'aprenentatge** (Barcelona, 18 de maig de 2012); y las ponencias y comunicaciones se encuentran en: <http://xtec.gencat.cat/ca/curriculum/xarxacb/jornades/jornada2012>

experiencias de éxito durante la implementación, de forma que la respuesta positiva de los estudiantes motive a los docentes a continuar (Sondergeld et al., 2010). En relación con la propuesta de sistema de evaluación que presenta el libro, resaltamos la idea de recoger todas las experiencias reales y compartirlas, como medida para motivar a la comunidad educativa en general (López-Pastor y Pérez-Pueyo, 2017). Por último, de las formaciones realizadas por la red de competencias básicas, destacamos la necesidad de trabajar en red con otros centros (Xarxa de Competències Bàsiques, 2019).

Después de analizar por qué los docentes deben dedicar tiempo y recursos en desarrollar su competencia evaluativa, justificando la necesidad de invertir en la alfabetización inicial y en la formación continua sobre evaluación para el aprendizaje, queremos presentar en el siguiente apartado nuestra propuesta. Ésta se basa en los lineamientos pedagógicos que hemos presentado como los más efectivos para conseguir la transferencia real de los aprendizajes realizados en la formación docente y el desarrollo sostenible de la innovación en evaluación. Además, como propuesta innovadora y basándonos en los aspectos descritos en el apartado sobre evaluación auténtica, la formación versará sobre cómo introducir la evaluación auténtica en el aula y que ésta sea un pilar del sistema de evaluación del centro. Pasamos pues a presentar lo que hemos denominado el Proyecto GAPPISA.

1.4.3. Formación docente sobre evaluación: Proyecto GAPPISA.

El Proyecto GAPPISA es un programa de formación continua para docentes de educación obligatoria, para los niveles de infantil, primaria y secundaria. La formación está centrada en ofrecer una base teórica y práctica sobre evaluación auténtica de los aprendizajes, con el propósito final de que los docentes puedan modificar sus exámenes de evaluación tradicionales, basadas en la valoración de contenidos, y diseñar pruebas de evaluación auténticas que promuevan el desarrollo de las competencias, el aprendizaje durante la evaluación y la atención a la diversidad.

El programa formativo, que se lleva a cabo a lo largo de un curso escolar con vistas a repetir el ciclo durante varios años, contempla actividades tanto de reflexión teórica como de acción en las que los docentes crean, mejoran y evalúan los resultados de los propios instrumentos diseñados. Esta estructura dota al proyecto de un carácter inminentemente práctico y contextualizado en la realidad docente de cada centro educativo, a la vez que promueve la sostenibilidad de este mediante la creación de una comunidad y cultura de evaluación.

En este apartado, después de una breve introducción sobre el origen del proyecto, explicaremos en qué se basa y qué consiste, detallando cuáles son sus principios formativos centrales y cómo se lleva a la práctica. Por último, describiremos las investigaciones que se han realizado basadas en esta formación y que dan cuenta de los resultados y su eficacia.

1.4.3.1. Origen del proyecto.

El nacimiento de esta formación se remonta al curso 2003, donde un instituto de educación secundaria obligatoria de la comarca de El Garraf, Catalunya, realizó una demanda de asesoramiento al equipo de investigación SINTE² especializado en estrategias de enseñanza y aprendizaje. La petición consistía en ayudar a los profesores a obtener mejores puntuaciones en las pruebas del Programa Internacional para la Evaluación del Alumnado (PISA). Para tal fin, además de la programación de sesiones de reflexión teórica sobre evaluación, los investigadores diseñaron un instrumento específico que sirviera para reconvertir los exámenes que normalmente utilizaban los profesores en pruebas de evaluación tipo PISA.

El instrumento diseñado resultó una guía para el análisis de pruebas, conformada por características surgidas de un proceso de extrapolación de algunos de los criterios utilizados en la construcción de los ítems de PISA. Esta se denominó Guía para el Análisis de las Pruebas desde la Perspectiva PISA, los docentes podían hacerse preguntas específicas sobre si sus pruebas cumplían o no los criterios

² Para más información sobre el equipo de investigación SINTE de la Universitat Autònoma de Barcelona, consultar el siguiente enlace: <https://www.sinte.me/>

necesarios para evaluar competencias de una forma realista y, a partir de aquí, introducir cambios en esta dirección.

El asesoramiento sobre evaluación con el uso de esta guía duró dos cursos escolares durante los cuáles participaron un total de ocho profesores. Sobre esta primera experiencia se realizaron varias investigaciones y publicaciones para dar cuenta de los resultados positivos de la formación y, especialmente, de la Guía GAPPISA como instrumento (Duran y Monereo, 2008; Monereo, Castelló et al., 2009; Monereo y Lemus, 2009;). Las mejoras significativas encontradas en los cambios en las pruebas de evaluación avivaron la implementación de la formación con el uso de la Guía GAPPISA en otros centros docentes, de forma que el proyecto se ha ido renovando y actualizando, con el uso de la Guía GAPPISA estableciéndose como eje vertebrador del mismo.

En la actualidad, la formación mantiene el mismo propósito de dotar de recursos y habilidades a los docentes para que estos mejoren sus pruebas, realizando cambios a partir de la realidad de cada centro y en dirección a las bases compartidas con las pruebas PISA. Sin embargo, a lo largo de muchas implementaciones, la guía GAPPISA ha evolucionado pasando de ocho a diez dimensiones, con el fin de transformarse en una evaluación realmente competencial, formativa e inclusiva. Y, además, a nivel de formación se han introducidos mejoras metodológicas para asegurar la eficacia de la formación y la sostenibilidad de las mejoras introducidas, surgidas a través de las necesidades detectadas en los diversos contextos donde se implementado. En los siguientes puntos detallaremos en qué consiste la guía y qué prácticas formativas se han utilizado, pero antes, es necesario profundizar un poco más en los lineamientos teóricos del proyecto.

1.4.3.2. ¿En qué se basa el Proyecto GAPPISA?

Como hemos adelantado, el propósito del Proyecto GAPPISA es conseguir que los docentes apliquen en sus aulas una evaluación de los aprendizajes auténtica, que posibilite el desarrollo de las competencias de los alumnos y, a la vez, mejorar su aprendizaje según las necesidades de cada uno. Esta idea de diseñar instrumentos

de evaluación auténticos parte, por un lado, de la investigación inicial que hemos destacado anteriormente y, por lo tanto, de las bases psicoeducativas implícitas en las pruebas PISA. Así pues, en este punto, analizaremos cuáles son estas bases, para en puntos posteriores, poder describir cuáles son las características que destacamos en el diseño de los instrumentos utilizados durante la formación. Por otro lado y, profundizando en parte una de las bases más destacadas de las pruebas PISA, la autenticidad, justificaremos por qué nos centraremos en la evaluación auténtica.

1.4.3.2.1. Importancia de las bases psicoeducativas de las pruebas PISA.

Hemos comentado anteriormente que la demanda de la que parte la formación era mejorar los resultados de las pruebas de evaluación del programa PISA, por lo tanto, el instrumento principal de la formación, la Guía GAPPISA, fue construida en relación con sus bases psicopedagógicas. Primero, es necesario destacar que las bases psicoeducativas se enmarcan en el enfoque competencial definido por la Comisión Europea en su marco de referencia y en el que se define la competencia como el conjunto de conocimientos, capacidades y actitudes adecuadas para dar respuesta a una situación o problema, combinándolos y poniéndolos en acción (Comisión Europea, 2007). De esta manera, las pruebas PISA, partiendo de la evaluación de competencias propone un conjunto de ítems a resolver, ordenados de menor a mayor complejidad cognitiva.

Todas las pruebas PISA, comparten un proceso de evaluación de la competencia caracterizado por tener: inicio con un estímulo previo, temas a tratar auténticos y significativos en relación con varios contextos, una creciente complejidad cognitiva, diversidad en la tipología de ítems referente tanto a los datos de la pregunta como las posibilidades de la respuesta y un formato de prueba según se responda con medios analógicos o digitales. En la Tabla 13, se puede observar un breve análisis de estas características en las pruebas PISA realizadas para cada una de las últimas competencias centrales, es decir, PISA 2009 con el foco en la Competencia Lectora, PISA 2012 con el foco en la Competencia Matemática y PISA 2015 con el foco en la Competencia Científica.

Tabla 13*Características de las pruebas PISA según año y competencia clave evaluada*

Característica	PISA 2009. Competencia Lectora	PISA 2012. Competencia Matemática	PISA 2015. Competencia Científica
Estímulo previo	Se inicia con una situación realista o un texto.	Se inicia con un problema realista.	Se inicia con un texto o problemática inicial, con imagen o no, que contextualiza
Autenticidad y significatividad	Situaciones de los ámbitos personal, público, educativo y laboral.	Situaciones de los ámbitos personal, educativo, social y científico.	Contextos personales, locales y nacionales y globales.
Complejidad cognitiva	Procesos de creciente complejidad, cinco acciones.	Procesos de creciente complejidad, cuatro acciones.	Tres niveles de demanda cognitiva de creciente complejidad.
Tipología de ítems según pregunta	Texto e imagen.	Texto e imagen (en la opción de soporte informático, algunas son interactivas)	Texto e imagen (algunas interactivas)
Tipología de ítems según respuesta	Cerrada y abierta escrita (frase a párrafo de cuatro) o dibujada (gráfico o diagrama)	Principalmente cerrada.	Cerrada y abierta escrita (frase a párrafo de cuatro) o dibujada (gráfico o diagrama)
Formato de la prueba	Papel	Papel; soporte informático opcional	Medios informáticos

Nota. Extraído de Mejías y Monereo (2018, p. 71)

Como podemos observar en la Tabla 13, no se dan grandes diferencias entre pruebas, a excepción del formato de la prueba y la tipología de algunos ítems, que se debe a la constante actualización por la que pasa el programa. De hecho, en este análisis, no aparecen el resto de las competencias paralelas que se incluyen en cada edición como áreas de innovación; estas han sido: lectura electrónica, resolución de problemas y resolución colaborativa de problemas. Estos nuevos dominios introducen dos nuevas características exclusivas como el uso estratégico de las TIC como instrumento y no solo como formato, y el aprendizaje colaborativo, aspectos ambos claves para una actuación competencial en el siglo XXI.

En resumen, las características que conforman las bases psicoeducativas de las pruebas PISA son el desarrollo de las competencias básicas, la autenticidad como

punto de partida para la propuesta de problemas complejos, la resolución de la situación auténtica planteada mediante diferentes tareas y una creciente complejidad cognitiva. De esta manera, las pruebas PISA realizan una evaluación competencial centrada en la resolución de un estímulo auténtico mediante la respuesta a diferentes ítems variados y ordenados de menor a mayor complejidad cognitiva.

En diversos estudios se ha demostrado que el análisis de los ítems de las pruebas PISA pueden ayudar a transformar la evaluación, incluso a que los docentes introduzcan modificaciones en su día a día de aula (Alcañiz y Cervera, 2014; Muñoz y Charro, 2017; Sanmartí y Sardà, 2007). En el Proyecto GAPPISA, partimos de esta premisa y los utilizamos para diseñar la Guía GAPPISA.

1.4.3.2.2. La evaluación auténtica como escenario motivador.

Por otro lado, otra de las características de la Guía GAPPISA es la elección de la evaluación auténtica como contenido central por varios motivos. En primer lugar, porque es una de las características más destacables de las bases psicoeducativas de las pruebas PISA. En segundo lugar, porque permite realizar una evaluación coherente con el enfoque competencial y la evaluación orientada al aprendizaje. Y, en tercer lugar, porque el término autenticidad es altamente atractivo en el mundo educativo, tanto para docentes como para alumnos, puesto que alude a la realidad y dota de relevancia los aprendizajes. Para justificar estos motivos, y aunque ya hemos desarrollado en profundidad en el apartado anterior lo que significa la evaluación auténtica, vamos a retomar algunas de sus características principales, con el fin de establecer una relación con las bases psicoeducativas explicadas anteriormente.

En nuestro marco teórico, con relación a la autenticidad, destacan tres características: descripción de una situación problemática contextualizada, propuestas de actividades auténticas y un desarrollo de estas actividades que va de menor a mayor complejidad cognitiva. Estas características están en correlato directo con las pruebas PISA en tanto que contienen, como hemos dicho anteriormente, un estímulo inicial auténtico y significatividad y complejidad cognitiva

en los diversos ítems. Además, el término autenticidad conecta con las acciones y motivaciones de los alumnos de forma intencional.

Sin embargo, nuestra propuesta teórica va un poco más allá y de la evaluación auténtica también destacamos su carácter competencial, su carácter inclusivo y su carácter formativo. Hablamos de que la evaluación consista en la resolución de una situación contextualizada donde destaca el realismo y la relevancia, y en la que para su resolución los alumnos deben usar activamente el material teórico y su diversidad de estrategias (Darling-hammond y Snyder, 2000; Gulikers, Bastiaens y Kirschner, 2006). También, de la importancia de la correspondencia de dicho realismo con cada una de las actividades de resolución del problema complejo (A. Herrington y J. Herrington 2006). Y, por último, de la necesidad de demostrar el proceso y de recibir retroalimentación mientras se realiza la evaluación (Ashford-Rowea et al., 2014). Pensamos que estos criterios promueven una implicación del alumnado y de los docentes.

Por otro lado, la conexión entre la competencialidad, el carácter inclusivo y el formativo es muy alta, lo que dota de consistencia teórica a la propuesta. Por ejemplo, si comprendemos que un problema auténtico es siempre complejo y que, para afrontarlo, se requiere ir más allá del conocimiento y aplicarlo en situaciones realistas, es necesario que las tareas de evaluación sean competenciales. También, la autenticidad nos remite a la idea de aprender haciendo, de mejorar a través de descubrir las posibilidades de actuación y de los fallos realizados, esto es, formarse mejor a través de la evaluación de las situaciones resueltas y, por lo tanto, que las actividades incluyan acciones que permitan la reflexión sobre el aprendizaje, que sean formativas. Por último, ambas características se relacionan directamente con la idea de que cada proceso de resolución de actividades realistas es único en tanto que la persona que lo afronta usa sus propios recursos, estrategias, maneras de investigar, de comunicar y de valorar su éxito, por lo tanto, las propuestas son inclusivas.

En definitiva, el conjunto de características que acabamos de apuntar y que en posteriores puntos definiremos en profundidad, pensamos que otorgan a la evaluación auténtica la cualidad de marco perfecto para la evaluación dentro del

nuevo enfoque de enseñanza por competencias y que, además, es altamente atractivo para los estudiantes en tanto que resulta un aprendizaje relevante y significativo para ellos.

1.4.3.3. ¿En qué consiste el Proyecto GAPPISA?

El Proyecto GAPPISA es más que una formación docente sobre evaluación puntual y descontextualizada. Uno de sus propósitos es conseguir que la formación alcance una transferencia real de los aprendizajes de los profesores, es decir, que después de la formación los docentes evalúen de forma auténtica en su contexto cotidiano de aula. También, que a nivel de centro educativo se construya una comunidad que aprende sobre evaluación y favorece alcanzar un consenso sobre cómo hacer real una evaluación auténtica. Para conseguir estos dos propósitos es importante diseñar una estructura formativa holística que permita el desarrollo de docentes e instituciones de forma efectiva y eficaz. En este punto vamos a explicar, en primer lugar, los principios formativos en los que se basa la formación y, en segundo lugar, cómo se llevan a la práctica, describiendo su estructura, planificación e instrumentos y recursos utilizados.

1.4.3.3.1. Características formativas.

En la descripción del proyecto GAPPISA a nivel pedagógico, vamos a retomar las características de calidad de las formaciones permanente del profesorado que se han demostrado como las más eficientes para conseguir mejorar los resultados de los alumnos, descritas en un apartado anterior. En la Tabla 14, exponemos, para cada característica, las acciones que se llevan a cabo en el Proyecto GAPPISA con las que se corresponden.

Tabla 14

Relación entre las características para una formación docente eficiente y las características del Proyecto GAPPISA

Características de las formaciones docentes eficientes	Acciones del Proyecto GAPPISA que aseguran la característica
Partir de la realidad del docente (sus prácticas, sus necesidades,	Los profesores mejoran un instrumento de evaluación que ya utilizan porque quieren mejorar su evaluación y que sea auténtica.

sus creencias, etc.) vinculando la formación a ésta.	
Permitir actividades en las que se pueda reflexionar sobre la propia práctica.	Los profesores participan en un seminario inicial en el que se recogen sus necesidades. Al final del proyecto se pide reflexionar sobre su práctica, concretamente, sobre la implementación de la evaluación auténtica diseñada.
Programar un ciclo formativo completo con fases claras e interconectadas.	La formación está estructurada de forma que las actividades van desde el aprendizaje de los contenidos necesarios, pasando por el diseño de los instrumentos, para llegar a la puesta en práctica de estos y finalizar con una reflexión sobre estos; las actividades dependen unas de las otras.
Diseñar actividades de aprendizaje basadas en la acción.	Los profesores diseñan e implementan un instrumento de evaluación ajustado a su realidad de aula con el cuál realizan un cambio real.
Programar tiempo real para poner en práctica los aprendizajes y acompañar esta aplicación.	La formación está programada para desarrollarse a lo largo de un curso escolar, de forma que los profesores dispongan de tiempo suficiente para reflexionar, crear, aplicar y evaluarse.
Disponer de un liderazgo pedagógico que vele por la construcción de una comunidad educativa que aprende y se co-responsabiliza.	La formación está coordinada por el líder pedagógico del centro, junto con los coordinadores de cada centro, capacitados para acompañar el aprendizaje de sus compañeros y promover la participación de nuevos profesores. Para ello, los coordinadores cuentan con una formación específica y la comunicación constante con los asesores.
Promover y asegurar la transferencia de los aprendizajes en el tiempo, a través de la autoevaluación de la implementación y la evaluación del impacto.	Los profesores realizan una última actividad en la que reflexionan sobre el impacto de la implementación de una evaluación auténtica. Además, cumplimentan un cuestionario de valoración de la formación con el que evaluamos su efectividad. Por otro lado, los coordinadores son los encargados de conseguir transformar el grupo de participantes en una comisión sobre evaluación estable que se dedicara a promover la participación de nuevos compañeros. Todos los instrumentos, además, están pensado para que los profesores puedan utilizarlos siempre que lo necesiten.
<i>Nota.</i> Las características formativas que exponemos y los estudios que las justifican están desarrolladas en el apartado Formación continua sobre evaluación.	

Como podemos ver en la Tabla 14, el proyecto se caracteriza por tener una programación de actividades anual que parten de la zona de desarrollo de los docentes y posibilitan n la acción-reflexión-acción y la creación de una comunidad evaluativa en el centro que perdure. Estas características, citadas anteriormente en la literatura trabajada (Robinson et al., 2009; Iranzo, 2011; Johnson et al., 2017; Kraft et al., 2018; López-Pastor y Pérez-Pueyo, 2017) hacen que el Proyecto

GAPPISA pueda cumplir con sus propósitos. En los siguientes puntos explicaremos cómo se concretan estas características en su programación y desarrollo.

Por último, cabe destacar que la formación se lleva a cabo bajo una modalidad semi-presencial, en la que se combinan actividades formativas presenciales de reflexión conjunta entre docentes, algunas con el acompañamiento del equipo asesor y actividades formativas individuales con el acompañamiento de un asesor especializado a través de una plataforma virtual. Este no ha sido en ningún caso un impedimento que afecte a la calidad del programa. Al respecto, estudios como el de Powell, Diamond, Burchinal y Koehler (2010) han concluido, después de comparar la misma formación en su modalidad presencial y virtual, que los dos formatos son efectivos para mejorar la práctica del profesorado.

1.4.3.3.2. Estructura y programación.

El Proyecto GAPPISA se estructura en cuatro fases desarrolladas a lo largo de un curso escolar, de forma que las diferentes actividades se distribuyen a lo largo de los tres semestres. En estas fases (ver Figura 2) se cuenta con todos los pasos necesarios para aplicar los aprendizajes y reflexionar sobre ellos. En la primera fase se realiza una formación teórico-práctica sobre evaluación auténtica; en la segunda fase, analizan sus pruebas y las reelaboran utilizando la Guía GAPPISA; en la tercera fase, se aplican la prueba diseñada en una situación real de aula; y, para acabar, en la cuarta fase, analizan y reflexionan sobre los resultados de dicha aplicación y de la formación.

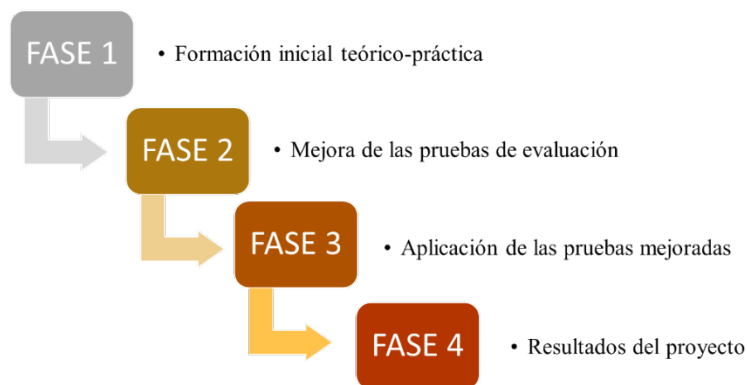


Figura 2. Esquema sobre las fases del Proyecto GAPPISA

A continuación, en las Tabla 15, exponemos la programación de cada fase para describir a fondo la propuesta formativa, para entender mejor la propuesta formativa, detallando las actividades, objetivos y recursos.

Tabla 15

Programación del Proyecto GAPPISA por fases y actividades

Fase 1. Formación inicial	
Actividad 1: <i>Seminario teórico-práctico</i>	
Previamente al seminario, los docentes reciben acceso a la Plataforma de formación y realizan la primera tarea que consiste en un cuestionario online sobre cómo es su evaluación. Estos resultados se utilizan para abrir el seminario presencial y realizar la primera reflexión conjunta sobre qué es la evaluación para los participantes. El seminario se divide en dos partes, una primera dedicada a explicar qué es la evaluación auténtica y las características de ésta a través de la Guía GAPPISA, y una segunda consistente en un taller práctico de análisis de casos, donde los profesores analizan pruebas con el fin de familiarizarse con las dimensiones de la Guía GAPPISA. Para acabar, se describe el proyecto, sus actividades y los recursos necesarios para realizarlo con éxito.	
Objetivo/s: <ul style="list-style-type: none"> • Reflexionar sobre la cultura de evaluación del centro y cómo es la evaluación que realizan los profesores. • Analizar las características de la evaluación auténtica. • Conocer las fases, actividades e instrumentos del Proyecto GAPPISA. 	
Recursos: dossier de trabajo y acceso a la plataforma ³ .	Modalidad: Presencial
Resultados: reflexión sobre la propia evaluación e identificación de qué aspectos pueden mejorarse.	Temporalización: 10h aprox. durante el primer trimestre.
Actividad 2: <i>Aprendizaje autónomo</i>	

³ La plataforma de aprendizaje se encuentra en la siguiente dirección:
<http://sintedi.wixsite.com/proyectogappisa>

En esta actividad, los profesores entran en la Plataforma virtual y realizan las diferentes tareas que se les propone (lectura, vídeo, análisis de casos, etc.) para profundizar en los conocimientos necesarios para comprender la evaluación auténtica.	
Objetivo/s: <ul style="list-style-type: none"> • Profundizar en las bases de la evaluación auténtica. • Saber cuáles son las dimensiones de la Guía GAPPISA y cómo podemos autenticar las pruebas de evaluación. 	
Recursos: acceso a la plataforma.	Modalidad: virtual.
Resultados: conocimientos sobre cómo autenticar una prueba de evaluación.	Temporalización: 3h aprox. durante el primer trimestre.
Fase 2. Elaboración de pruebas auténticas	
Actividad 3: Autenticación de la prueba de evaluación	
Los profesores deben analizar una prueba de evaluación que utilicen habitualmente durante el segundo o tercer trimestre de su docencia, con la ayuda de la Guía GAPPISA y autenticarla, es decir, realizar cambios en la prueba para conseguir que sea más auténtica. Después, deben enviar la prueba original y la versión autenticada a los asesores para recibir feedback.	
Objetivo/s: <ul style="list-style-type: none"> • Saber utilizar la Guía GAPPISA para autenticar pruebas de evaluación. • Diseñar una prueba de evaluación auténtica. 	
Recursos: acceso a la plataforma, la Guía GAPPISA (ver Anexo 1) y una prueba de evaluación propia.	Modalidad: virtual.
Resultados: la versión autenticada de la prueba de evaluación.	Temporalización: 5h aprox. durante el segundo trimestre.
Actividad 4: Reunión de seguimiento entre asesores y coordinadores	
Los asesores se reúnen virtualmente con los coordinadores de centro para compartir los progresos de los profesores y los primeros resultados.	
Objetivo/s: <ul style="list-style-type: none"> • Comunicar a los coordinadores de centro los resultados de las primeras actividades. • Establecer las medidas de apoyo necesarias. 	
Recursos: acceso a la plataforma.	Modalidad: virtual.
Resultados: avance de los profesores y establecimiento de estrategias de mejora.	Temporalización: 1h aprox. durante el segundo trimestre.
Actividad 5: Reunión de seguimiento del coordinador con su equipo de profesores	
Los coordinadores se reúnen con su equipo de profesores para reflexionar sobre los avances y los cambios a realizar antes de entregar la última versión de la prueba de evaluación auténtica.	
Objetivo/s: <ul style="list-style-type: none"> • Compartir las dificultades y los avances para seguir mejorando las pruebas. 	
Recursos: acceso a la plataforma.	Modalidad: presencial.
Resultados: mejora de la implicación de los participantes y de sus productos.	Temporalización: 1h aprox. durante el segundo trimestre.
Actividad 6: Mejora de la prueba auténtica	
Previamente, los profesores han recibido un feedback de la prueba autenticada que diseñaron, que incluye una rúbrica donde se detallan los cambios que se pueden hacer en cada dimensión de la Guía GAPPISA. Con este feedback los profesores deben mejorar su prueba y elaborar la versión final que aplicarán en la siguiente fase. Para acabar, deben enviarla de nuevo para poder recibir el último feedback.	
Objetivo/s:	

<ul style="list-style-type: none"> • Saber utilizar la Rúbrica de Evaluación de las Pruebas Auténticas. • Mejorar la prueba de evaluación auténtica diseñada mediante el feedback recibido por el equipo de asesores. 	
Recursos: acceso a la plataforma, el feedback del equipo de asesores y la Rúbrica de Evaluación de Pruebas Auténticas (ver Anexo 2).	Modalidad: virtual.
Resultados: la versión final de la prueba de evaluación auténtica.	Temporalización: 3h aprox. durante el segundo trimestre.
Fase 3. Aplicación de las pruebas auténticas	
Actividad 7: Aplicación de la prueba de evaluación en una situación real	
Los profesores deben utilizar la prueba de evaluación auténtica que han diseñado en su clase y recoger evidencias sobre los cambios que se han producido en la enseñanza y en el aprendizaje de los alumnos antes, durante y después de su aplicación. Para ello disponen de las preguntas del informe de reflexión que deberán completar en la siguiente actividad.	
Objetivo/s: <ul style="list-style-type: none"> • Aplicar una prueba de evaluación auténtica en el contexto de aula real. • Recoger evidencias de los resultados del uso de la prueba de evaluación auténtica. 	
Recursos: condiciones necesarias para aplicar la prueba auténtica y acceso a la plataforma.	Modalidad: presencial.
Resultados: la versión final de la prueba de evaluación.	Temporalización: 3h aprox. durante el segundo o tercer trimestre.
Actividad 8: Reflexión sobre la aplicación y sobre la formación	
Después de la aplicación de la prueba, los profesores deben completar el Informe final de reflexión sobre cómo ha ido y qué efectos ha tenido la aplicación de la prueba de evaluación auténtica. Seguidamente lo envían a los asesores para que puedan analizar el impacto de la formación. También cumplimentan un cuestionario sobre la formación.	
Objetivo/s: <ul style="list-style-type: none"> • Analizar las evidencias recogidas antes, durante y después de la aplicación de la prueba auténtica • Reflexionar sobre el impacto de la aplicación de la prueba de evaluación auténtica. 	
Recursos: el Informe Final sobre la aplicación de la prueba (ver Anexo 3) y el cuestionario sobre la formación.	Modalidad: virtual.
Resultados: la versión final de la prueba de evaluación auténtica.	Temporalización: 2h aprox. durante el segundo o tercer trimestre.
Fase 4. Evaluación de los resultados y de la formación	
Actividad 9: Reunión de resultados entre asesores y coordinadores	
Los asesores se reúnen con los coordinadores y con el equipo de liderazgo del centro para comunicar los resultados globales de la formación: participación, nivel de las pruebas diseñadas, resultados de la aplicación de esta y valoración final de la formación.	
Objetivo/s: <ul style="list-style-type: none"> • Comunicar los resultados de la formación y realizar una valoración conjunta. 	
Recursos: acceso a la plataforma.	Modalidad: virtual.
Resultados: Informe Final sobre la Formación.	Temporalización: 1h aprox. durante el tercer trimestre.
Actividad 10: Comunicación de los resultados y constitución de la comisión sobre evaluación	
Los coordinadores se reúnen con su equipo de profesores para comunicarles los resultados de la formación analizados por el equipo asesor y para constituir una comisión sobre	

evaluación encargada de construir una cultura evaluativa en el centro que promueva la participación de nuevos profesores en la formación.	
Objetivo/s: <ul style="list-style-type: none"> • Reflexionar sobre el impacto que la formación ha tenido en los profesores y en sus prácticas de evaluación. • Constituir una comisión de evaluación encargada de promover la implementación de la innovación evaluativa en el centro educativo. • Definir los objetivos de la comisión de evaluación para el próximo curso. 	
Recursos: acceso a la plataforma y a la presentación de los resultados del informe final.	Modalidad: presencial.
Resultados: comisión sobre evaluación.	Temporalización: 1h aprox. durante el tercer trimestre.
Actividad 11: Acceso a los recursos GAPPISA para continuar autenticando pruebas	
Toda la comunidad educativa que ha participado en el Proyecto GAPPISA tendrá acceso a los instrumentos necesarios para seguir autenticando sus pruebas de evaluación (la Guía GAPPISA y su Rúbrica). También tendrán acceso a un nuevo recurso elaborado por los asesores: el banco de pruebas auténticas, en el que se recogen las pruebas de calidad diseñadas por los docentes de todos los centros participantes.	
Objetivo/s: <ul style="list-style-type: none"> • Disponer de recursos para continuar autenticando pruebas de evaluación • Promover la sostenibilidad de los cambios a través del acceso a los recursos. 	
Recursos: acceso a la plataforma y a la base de recursos.	Modalidad: virtual.
Resultados: página web con todos los instrumentos y la base de datos de pruebas auténticas ⁴ .	Temporalización: variable.

Esta programación cumple los criterios necesarios para conseguir que los participantes se impliquen en la formación, en tanto que disponen de tiempo e instrumentos específicos para trabajar en sus cambios, pueden reflexionar sobre su aprendizaje y los resultados que este tiene y establecen nuevos objetivos para continuar con la innovación en evaluación. A continuación, explicaremos con más detalle los instrumentos y recursos de los que disponen.

1.4.3.3.3. Instrumentos y recursos.

Debido a que la formación es semipresencial y a su carácter de aprendizaje activo en el que los docentes aprenden haciendo, el Proyecto GAPPISA cuenta con un amplio abanico de instrumentos formativos y de recursos para hacer efectivo el

⁴ La plataforma que recoge todo los recursos sobre evaluación auténtica del proyecto GAPPISA es la siguiente: <https://sintedinetwork.wixsite.com/evaluacionautentica>

aprendizaje de los docentes. Por un lado, como vemos en la Tabla 16, encontramos recursos que permiten la participación de los profesores en la formación y promueven y facilitan la comunicación con y entre éstos. Por otro lado, disponemos de un conjunto de instrumentos que son posibilitadores y mediadores del aprendizaje de los contenidos y del cumplimiento de los objetivos de cada actividad.

Tabla 16

Descripción de los recursos y los instrumentos formativos del Proyecto GAPPISA

Nombre	Descripción	Acceso y/o consulta
Recurso: <i>Plataforma de aprendizaje</i> del Proyecto GAPPISA.	Plataforma virtual de la formación que contiene actividades e instrumentos para el aprendizaje.	Dirección de la web: https://sintedi.wixsite.com/proyectogappisa
Recurso: Banco de recursos de Evaluación Auténtica	Página web con acceso a la Guía GAPPISA, a la webquest del proyecto y al Banco de pruebas auténticas.	Dirección de la web: https://sintedinetwork.wixsite.com/evaluacionautentica
Instrumento: Guía GAPPISA.	Instrumento Guía GAPPISA que contiene preguntas de reflexión y criterios a seguir para autenticar una prueba de evaluación tradicional.	Formato en papel centrado en la auto-reflexión sobre la propia prueba: https://docs.wixstatic.com/ugd/8ef2e0_8b6fb0c592c84cc19a5e53ee10fe63b5.pdf También disponible en el Anexo 1. Formato online centrado en el diseño de a través de ejemplos: https://sintedinetwork.wixsite.com/gappisa
Instrumento: Rúbrica de evaluación de las pruebas auténticas.	Instrumento para evaluar las pruebas autenticadas en el que se muestran los diferentes indicadores de calidad para diseñar una prueba de calidad.	Formato online de acceso exclusivo para los participantes. Disponible en el Anexo 2.
Instrumento: Informe Final sobre los Resultados de la aplicación	Instrumento para analizar los cambios producidos antes, durante y después por la aplicación de la prueba de evaluación auténtica.	Formato online de acceso exclusivo para los participantes. Disponible en el Anexo 3.

En primer lugar, tenemos la *Webquest* del Proyecto GAPPISA, la creación de la cual respondió, en un primer momento, a la necesidad de disponer de una plataforma donde poner a disposición de los participantes los materiales necesarios para realizar la formación, debido especialmente a su carácter semipresencial. Sin embargo, finalmente optamos por cumplir al máximo las características de las *webquest* como estrategia didáctica basada en los recursos en línea (Adell,

Mengual-Andrés y Roig-Vila, 2015); de esta manera la plataforma ofrece una estructura del proyecto para guiar a los docentes, acceso a diferentes recursos tanto escritos como audiovisuales, consulta directa mediante cuestionarios abiertos, interacción en las redes sociales para socializar el aprendizaje, videoconferencias programadas, etc., convirtiéndose en un entorno virtual de aprendizaje.

Respecto a las actividades, es preciso indicar que están agrupadas siguiendo las fases del desarrollo del proyecto y diferenciando entre obligatorias y de ampliación, tareas que promovían una mayor implicación y profundización en los contenidos. Además, la redacción de las propuestas de actividades es positiva y anima a los docentes a continuar, celebrando sus avances y promoviendo reflexiones.

Y, en relación con la socialización del aprendizaje, se enfocó al objetivo de construir una comunidad sobre evaluación más allá de la formación y conectando con otras redes, en la que poder compartir ideas y experiencias. Para ello, en la plataforma, algunas tareas proponen compartir respuestas a preguntas y otras ideas en las redes sociales, específicamente en Twitter a través de la etiqueta *#proyectogappisa* y *#evaluaciónauténtica*, y en Facebook dentro de la página Evaluación Auténtica - Sintedi⁵, donde además los asesores comparten información y noticias de interés sobre la evaluación.

En segundo lugar, el Banco de Recursos de Evaluación Auténtica, responde a la necesidad de diseñar un Banco de Pruebas Auténticas, en el que se recogen las mejores pruebas de evaluación diseñadas por los docentes a lo largo de las diferentes ediciones del Proyecto GAPPISA. Los docentes que han participado pueden acceder, buscar pruebas según el nivel educativo o las áreas trabajadas, leer una ficha breve con su descripción donde se indica la autoría y descargarse la prueba. El acceso a estas pruebas de calidad no está pensado para que los docentes se descarguen la prueba y la apliquen tal cual, en su aula, sino para que puedan analizar el trabajo de otros compañeros y coger ideas para diseñar sus

⁵ La dirección es la siguiente: <https://www.facebook.com/projectegappisa/>

propias pruebas. Además, desde el banco de recursos se facilita el acceso a la Guía GAPPISA y a la Plataforma virtual. Por último, aunque todavía en desarrollo, se puede acceder a un apartado donde se explican experiencias reales de profesores que realizan una evaluación auténtica.

En tercer lugar, la Guía GAPPISA, en la que profundizaremos en apartados siguientes, es un instrumento de auto-reflexión para analizar los niveles de competencialidad, inclusividad y poder formativo de las pruebas de evaluación. Consta de diez dimensiones que definen una evaluación auténtica y un conjunto de preguntas que dirigen la atención a los aspectos que se deberían modificar. Así pues, sirve como instrumento teórico, mediante el cual aprender qué características conforman una evaluación auténtica, y como instrumento práctico, en tanto que los docentes introducen cambios a partir de esta reflexión. Además, la guía va acompañada con un documento explicativo en el que se detalla por qué es importante cada dimensión y qué cambios pueden hacerse para conseguir una evaluación competencial, formativa e inclusiva. En su expresión virtual, se incluyen además algunos ejemplos reales para cada dimensión.

En cuarto lugar, la Rúbrica de evaluación de las Pruebas Auténticas permite a los docentes tener un instrumento más para analizar sus pruebas. A diferencia de la Guía GAPPISA, la rúbrica permite conocer cuál es el siguiente paso para mejorar la prueba, es decir, detalla cuáles son los criterios de calidad de cada dimensión. También se recogen en esta Rúbrica cuáles son los cambios que pueden realizar los docentes en las pruebas pero que no los conducen a una mejora, sino a lo que hemos denominado una *pseudoautenticación*. Estos cambios nacen del análisis de las pruebas de evaluación de docentes que realizaron una incorrecta interpretación, y que suponen una oportunidad de aprendizaje para el resto.

En quinto lugar, el Informe Final sobre los resultados de la aplicación de la prueba, plantea un conjunto de preguntas para reflexionar sobre los efectos que la prueba ha tenido en el aprendizaje de los alumnos y en la manera de enseñar de los docentes, agrupadas en tres momentos: uno, relativo a antes de la prueba, con aspectos referentes al estudio de los alumnos y las actividades previas a la prueba; dos, relativo a durante la prueba, con aspectos como el uso de recursos, previsión

del tiempo, tipo de ayudas, etc.; y tres, relativos a después de la prueba, con aspectos como los resultados de la prueba, incidencias o quejas generadas, etc. Este informe, también permite preparar la aplicación de la prueba.

La Guía GAPPISA

Como indicamos anteriormente, la Guía GAPPISA es un instrumento que permite reflexionar sobre los aspectos que hacen que una prueba sea auténtica, en tanto que posibilita la evaluación de las competencias promoviendo el aprendizaje a través de ésta y ajustándose a las necesidades de los alumnos. Elegimos la guía como instrumento porque posibilita que el profesor reflexione y sepa qué preguntas realizarse para mejorar sus instrumentos de evaluación. En este apartado vamos a profundizar en el contenido de esta, explicando sus categorías, su relación con las pruebas PISA y las dimensiones que la conforman.

La Guía GAPPISA está compuesta por diez características que, desarrolladas en mayor o menor grado, determinan la autenticidad de la prueba o instrumento de evaluación. Estas características se agrupan en tres bloques principales de la siguiente manera:

- Categoría de competencialidad en la que encontramos: situación problema inicial, actividades auténticas, actividades de diferente complejidad cognitiva y condiciones de resolución de la prueba auténticas.
- Categoría de potencial formativo en la que encontramos: evaluación del proceso, puntuaciones y criterios compartidos y comunicación de los resultados.
- Categoría de inclusividad en la que encontramos: formato heterogéneo de los enunciados, formato heterogéneo de las respuestas y atención a la diversidad.

Esta agrupación por categorías es únicamente orientativa y se utiliza a efectos de poder determinar sobre qué categoría se identifican mayores cambios. Es decir, hay una relación de interdependencia entre todas las características, de forma que modificar una promueve un nuevo cambio, como podremos apreciar cuando describamos cada característica.

Pero, antes de profundizar, es preciso señalar que la propuesta que presentamos actualmente difiere de la primera edición (Duran y Monereo, 2008), en la que solo contaba con siete dimensiones y más dirigidas al aspecto competencial. La Guía GAPPISA se ha actualizado para incluir nuevos aspectos teóricos relacionados con la atención a la diversidad y la evaluación para el aprendizaje. Además, algunas de las dimensiones se han hecho más explícitas y descrito mejor, facilitando al docente la comprensión de estas. Estos cambios han surgido de la experiencia de la propia formación, a través de la cual se ha nutrido la guía. De esta forma la guía GAPPISA responde a las nuevas realidades y necesidades formativas sobre evaluación del profesorado.

Sin embargo, en ningún momento han dejado de estar presentes las características iniciales ni de ser evidente su relación con los orígenes de esta que hemos tratado en apartado anteriores, las pruebas PISA. Como exponemos en la Tabla 17, la mayoría de las categorías son coincidentes.

Tabla 17

Comparación entre las características de la Guía GAPPISA y las pruebas PISA

Características pruebas GAPPISA	Categorías centrales	Características pruebas PISA
Situación problema inicial contextualizada	Competencialidad	Estímulo previo
Actividades auténticas	Competencialidad	Autenticidad y significatividad
Complejidad cognitiva	Competencialidad	Complejidad cognitiva
Formato heterogéneo de los enunciados	Inclusividad	Tipología de ítems según pregunta
Formato heterogéneo de las respuestas	Inclusividad	Tipología de ítems según respuesta
Condiciones de resolución auténticas	Competencialidad	Formato de la prueba
Evaluación del proceso	Poder formativo	No se contempla.
Criterios de evaluación compartidos	Poder formativo	No se contempla.
Comunicación formativa de los resultados	Poder formativo	No se contempla.
Atención a la diversidad	Inclusividad	No se contempla.

Nota. Extraído de Mejías y Monereo (2018, p. 77)

Entrando en detalle, la categoría de competencialidad, a la cual nos referimos en la guía GAPPISA principalmente con las características de autenticidad y complejidad

cognitiva, es la que más peso tiene en el diseño de pruebas tipo PISA, en las cuales todos los problemas hacen referencia a contextos auténticos y su resolución se concreta teniendo que resolver actividades de diferentes niveles de dificultad. Respecto a la categoría de inclusividad, esta también encuentra una correlación entre la guía GAPPISA y las pruebas PISA, cuando nos fijamos en la diversidad de los ítems, sin embargo, al ser las PISA pruebas estandarizadas, no están diseñadas para atender ajustar más a las necesidades de los estudiantes. Por último, en relación con la categoría evaluación formativa, esta no encuentra un correlato directo con las pruebas PISA, siendo una concepción no aplicable por el mismo motivo anterior; aunque el hecho de resolver una situación a través de los diferentes ítems podría considerarse una manera de aprender a resolver el problema.

Las diferencias entre la Guía GAPPISA y las pruebas PISA se justifican claramente en cuanto recordamos los objetivos de cada una: mientras que las pruebas diseñadas con la Guía GAPPISA están pensadas para ser aplicadas a un grupo específico de estudiantes en el contexto de aula de un centro específico, las pruebas PISA deben poder aplicarse a una muestra significativa de la población de educación secundaria de un país determinado. Al fin y al cabo, el objetivo de PISA es analizar el sistema educativo de cada país y compararlo con otros, no evaluar y acompañar la mejora de los aprendizajes de los estudiantes.

Pasamos ahora a presentar con detalle la Guía GAPPISA. En la Tabla 18 describimos cada una de las características que hacen posible una evaluación auténtica, indicando también cuáles son las preguntas de reflexión que deben hacerse los profesores para conseguir que sus pruebas sean más competenciales, formativas e inclusivas.

Tabla 18
Descripción de las características de la Guía GAPPISA

Característica	Preguntas de reflexión	Descripción
1. Situación-problema inicial	¿Las distintas actividades de la prueba se enmarcan en una situación problema inicial debidamente contextualizada?	La prueba se inicia con la descripción de un problema a partir del cual se integran las distintas actividades y del que se describe de forma detallada el contexto físico y social.
2. Actividades auténticas	¿Las actividades de la prueba son realistas?	El problema puede encontrarse de forma similar en el contexto extra-escolar o

	<p>¿Las actividades de la prueba son relevantes?</p> <p>¿Las actividades de la prueba son socializadoras?</p>	<p>futuro académico-profesional. Se vincula de forma explícita la utilidad de la prueba para resolver situaciones familiares para el alumno. Las actividades permiten al alumno desarrollar su identidad ciudadana, académica y/o profesional.</p>
3. Actividades de diferente complejidad cognitiva	<p>¿Existen actividades de baja, media y alta complejidad cognitiva?</p> <p>¿Los niveles de complejidad cognitiva son progresivamente más elevados a lo largo de la prueba?</p>	<p>Las actividades de la prueba se relacionan con los procesos que definen una actuación competente: saber analizar la demanda, analizar las condiciones del contexto, definir un objetivo, planificar una resolución, utilizar los recursos disponibles, resolver el problema, evaluar el resultado, evaluar el proceso de resolución, comunicar y justificar tanto el resultado como el proceso seguido (por escrito u oralmente, pero acorde al discurso y género del ámbito), y ser capaz de aprender de la experiencia para resolver de modo más estratégico y eficiente demandas similares en el futuro. Se guía al alumno en el proceso de resolución de los problemas: las primeras actividades son de bajo nivel cognitivo y progresivamente se requieren tareas de nivel superior.</p>
4. Formato heterogéneo de los enunciados	<p>¿Las preguntas se presentan en diversidad de formatos (<i>in-put</i>)?</p>	<p>El formato de las distintas preguntas y de la información que se presenta en la prueba es suficientemente heterogéneo (textual, visual, auditivo...) El formato de <i>in-put</i> se corresponde con el que se daría en una situación real.</p>
5. Formato heterogéneo de las respuestas	<p>¿La prueba requiere respuestas del alumno en varios formatos (<i>out-put</i>)?</p>	<p>El formato de las distintas respuestas y las evidencias que se entregan es suficientemente heterogéneo (textual, visual, auditivo...) Las respuestas se presentan en el formato en el que se presentaría en una situación real.</p>
6. Condiciones de resolución de la prueba auténticas	<p>¿Las condiciones de resolución de la prueba son similares a las de la vida real?</p>	<p>Las condiciones de resolución (tiempo, recursos y materiales, personas implicadas e interacción entre ellas) son realistas.</p>
7. Evaluación del proceso	<p>¿La prueba requiere que el alumno explicita el proceso de resolución en algún momento?</p> <p>¿Se evalúa el ajuste del alumno a situaciones nuevas y/o inesperadas?</p>	<p>Para fomentar las estrategias de autorregulación de los alumnos, la prueba debe requerir que el alumno explicita (o muestre) el proceso de resolución que ha seguido en los tres momentos (antes, durante y después); de este modo lo podremos ayudar/guian mejor su aprendizaje. Se propician</p>

	¿Se evalúa el uso que el alumno hace de los recursos y/o ayudas?	situaciones nuevas e inesperadas en las que el alumno debe ajustarse “in situ” y se evalúa el modo en que ha regulado su respuesta. Se intenta “ver” y evaluar el modo en que el alumno utiliza las ayudas o recursos de consulta.
8. Puntuaciones y criterios compartidos	<p>¿Se comparte con el alumno el valor de cada pregunta/actividad?</p> <p>¿Se comparten con el alumno los criterios de corrección cualitativos?</p> <p>¿Se fomenta de forma eficiente la interiorización de los criterios de calidad?</p>	<p>Se comparten con el alumno los criterios de corrección cualitativos, además de cuantitativos, de cada pregunta para que el alumno pueda regular su actuación de forma estratégica y aprender durante la prueba. Para ello se describen las variables que harán que una respuesta sea “mala/regular/buena/muy buena” de modo que el alumno pueda interiorizarlos progresivamente. Además, para asegurar que están interiorizando dichos criterios, se realizan actividades de autoevaluación (el alumno corrige su propia prueba) y coevaluación (el alumno corrige la prueba de un compañero).</p>
9. Comunicación de los resultados	<p>¿El tipo de comunicación es bidireccional y cualitativa?</p> <p>¿La finalidad última de la comunicación es valorar el progreso intraindividual del alumno? ¿Se contextualiza el resultado del alumno con su progreso en pruebas anteriores?</p> <p>¿Las evidencias que se aportan al alumno son suficientes para fomentar su aprendizaje? ¿El tipo de correcciones hacen referencia a la calidad del proceso y respuesta correctos?</p> <p>¿Se posibilita al alumno revisar su prueba y construir nuevo aprendizaje basado en las correcciones hechas?</p>	<p>Con el fin de que la actividad de evaluación tenga sentido y suponga un aprendizaje en el que pueden profundizar más allá de la prueba, la comunicación de los resultados supone una parte esencial. Se recogerán las opiniones de los alumnos, respecto a los resultados obtenidos y los procesos realizados. A su vez, las correcciones que se retornen serán cualitativas y señalando el progreso de cada alumno respecto a pruebas anteriores. Para posibilitar el aprendizaje, se dará la opción al alumno de poder revisar la prueba con las correcciones realizadas.</p>
10. Atención a la diversidad	<p>¿Se han analizado los distintos perfiles de aprendizaje del grupo clase?</p> <p>¿Se han adaptado las características de la prueba para aquellos alumnos con necesidades educativas especiales (p. e. contenido, condiciones de resolución, tipo de corrección, etc.)</p>	<p>Para conseguir que todos los alumnos puedan participar y comunicar todos sus aprendizajes, se adaptará la prueba en base a los perfiles de aprendizaje y las necesidades educativas especiales. Pueden realizarse adaptaciones en relación con las ayudas que se dan, los recursos que se pueden emplear, el formato de la prueba, el tiempo de resolución, la profundización en el</p>

Las características expuestas en la Tabla 18, están ordenadas según la su interdependencia y considerando los cambios que los docentes pueden realizar con mayor facilidad, de forma que este orden guie al profesor a construir su evaluación auténtica. Así es, no podemos, por ejemplo, pensar en la evaluación del proceso, si no hemos planteado una situación suficientemente compleja que implique una sucesión de pasos de resolución en los que podamos observar un proceso de resolución. Por otro lado, el hecho de incluir una característica final que aluda a la atención a la diversidad de todos los alumnos, entendida como una respuesta a todos los perfiles de aprendizaje, responde a la complejidad de abordar este tema y la dificultad que tienen los docentes para hacerlo. Por último, observamos que las preguntas de reflexión permiten pensar en diferentes aspectos en los que puede manifestarse la característica, de forma que los docentes pueden desarrollarlas en mayor o menor medida, atendiendo a sus necesidades y habilidades.

1.5. La Identidad Docente y el Posicionamiento sobre Evaluación

Acabamos de presentar la propuesta de formación sobre evaluación auténtica justificando cómo cumplir su doble propósito: conseguir que los docentes utilicen la evaluación auténtica y que desarrollen su competencia en evaluación. Respecto al desarrollo de su competencia, y con el objetivo de que las mejoras evaluativas que vayan a introducir sean sostenibles en el tiempo, vemos necesario que esta formación produzca un cambio en la identidad del docente, específicamente, en su posicionamiento sobre evaluación.

Actualmente, la literatura sobre identidad docente ha aumentado considerablemente, de forma que la identidad es un constructo bien definido que puede emplearse en la investigación con muchos y diferentes propósitos. Uno de ellos es la comprensión del comportamiento y de los pensamientos de los profesionales de la educación para ayudarles en su desarrollo. Es decir, la identidad docente nos ayuda a entender cómo los profesores viven su profesión y cuáles son las representaciones que tienen sobre esto. También nos permite conocer cómo las experiencias en las que se ven implicados modifican sus ideas sobre qué significa ser y actuar como docente. Pero, en primer lugar, es importante concretar desde qué perspectiva nos posicionamos para definirla.

Desde la psicología de la educación de carácter más constructivista (Monereo y Badia, 2011) y desde la teoría dialógica (Hermans, 2001, 2009, 2014) la identidad del docente ha sido caracterizada como el conjunto de representaciones que tiene el profesional sobre "ser docente" relativamente estables en el tiempo (Badia et al., 2017) y como un proceso de (re)construcción durante la relación entre la persona y el contexto, en el que se manifiestan diferentes posiciones-del-yo mediante las cuales los sujetos actúan como agentes activos en el medio que les rodea (Beijaard, Meijer y Verloop, 2004).

Así pues, entendiendo la identidad desde el campo de la intersubjetividad, podemos destacar que el profesor no se caracteriza por tomar posición frente a la realidad de forma fija, estática e inamovible, sino más bien por actuar como un profesional que a medida que interacciona en un entorno educativo propio y compartido va

transformando sus concepciones, acciones y motivos. Desde los enfoques en los que nos situamos, por lo tanto, entre la dimensión intrapsicológica y la interpsicológica de la identidad se produce una construcción dinámica del posicionamiento del sujeto, cuando dialoga consigo mismo y con los otros. De esta manera, la identidad docente se caracteriza por ser estable y subjetiva, a la vez que múltiple, discontinua y social (Akkerman & Meijer, 2011).

Esto tiene fuertes implicaciones en la práctica en tanto que entender la identidad como un constructo dinámico puede condicionar la manera que los profesionales tienen de representarse sus roles y prácticas situadas. A partir de estas implicaciones, los expertos en la temática han diseñado formaciones que tienen como propósito acompañar a los docentes en sus procesos de desarrollo y/o actualización de sus competencias profesionales, y dotarlos de estrategias para superar los incidentes críticos que puedan surgir o hayan surgido durante la práctica, hecho que favorece la sostenibilidad de las innovaciones introducidas (Del Maestro y Monereo, 2014; Fernández y Marrero, 2013; Monereo, Weise et al., 2013; Weise y Sánchez-Busqués, 2013).

En suma, en los siguientes apartados, en primer lugar, examinaremos el porqué de las actuaciones de los docentes, es decir, cómo justifican sus decisiones, en qué creencias se apoyan, qué sentimientos las acompañan y si esto se traduce en una práctica coherente con los criterios que sostienen, para comprender cómo se define la identidad docente. En segundo lugar, a partir de este análisis identificaremos y describiremos de qué manera se posicionan los docentes sobre la evaluación y cuáles son los elementos que pueden determinarla. Y, en tercer lugar, propondremos dispositivos que permitan una reconstrucción de estas posiciones para que se alineen con el enfoque de evaluación trabajado en este estudio, facilitando el desarrollo de la competencia en evaluación de los docentes.

1.5.1. La Identidad Docente.

En este punto del apartado vamos a describir qué entendemos por identidad docente y qué aspectos son esenciales para interpretar cómo los docentes se posicionan respecto a la evaluación.

El primer aspecto de la identidad docente que debemos definir es su dualidad constitutiva, es decir, que la identidad está conformada por dos dimensiones según las cuáles la identidad es estática e individual y, a la vez, dinámica y social, erigiéndose una interrelación entre ambas que permite que el docente pueda ajustarse al contexto sin perder su esencia. La dimensión determinista se caracteriza por ser estable y continua en el tiempo, y la dimensión relativista se caracteriza por ser cambiante y constantemente construida socialmente (Akkerman y Meijer, 2010; Badia y Monereo, s.f.; Monereo, Weise et al., 2013). Entre ambas existe coherencia y un diálogo, de forma que la estabilidad de una posibilita la construcción situacional de la otra (Salgado y Hermans, 2005).

Así pues, por un lado, podemos definir la identidad docente como el conjunto de representaciones que tiene el profesor sobre su "ser docente" con relación a tres aspectos interrelacionados entre sí: las representaciones sobre su rol profesional, las representaciones sobre los enfoques que debe llevar a la práctica y las representaciones sobre los atributos afectivos asociados (Monereo y Pozo, 2011). El primer conjunto de representaciones, el rol profesional, hacen referencia a las concepciones que el docente tiene sobre la función y forma de la enseñanza y el nivel particular en el que actúa. El segundo grupo, los enfoques prácticos, da cuenta del conocimiento estratégico y didáctico especialmente de su disciplina y los métodos en los que se basa para diseñar y aplicar prácticas educativas durante su acción como docente. Y el tercer conjunto, los atributos afectivos, se dirige a los sentimientos, a los calificativos afectivos que acompañan lo anterior y la gestión que el docente realiza de aquello que siente.

Estas representaciones declaradas por el docente, por tanto, constituyen lo que podemos denominar su identidad narrada. La interrelación y coherencia entre las tres dimensiones es la que constituye una definición del sujeto como profesional de la docencia. En la literatura podemos encontrar diversos estudios sobre cómo esos componentes de la identidad se relacionan entre sí desde la voz de los docentes, tanto en contextos de docencia universitaria presencial (Monereo y Domínguez, 2014; Monereo, Weise et al., 2013) como en línea (Badia et al., 2017).

Por otro lado, cuando estas representaciones se ponen en acción constituyen lo que Monereo y Pozo (2011) en su origen denominaron *selfs*. En la activación del conjunto de representaciones plural se configuran *selfs* que pueden resultar más o menos efectivos, pero que siempre son afectivos, en tanto que afectan de manera implícita o explícita y condicionan la actuación. Además, no siempre la práctica, coincide con el conjunto de representaciones con las que el profesor se define. Por ejemplo, un profesor que afirma con firmeza que su evaluación está orientada al aprendizaje de los alumnos, frente a un grupo de alumnos puede escoger aplicar test teóricos y nada más, actuando desde una concepción de la evaluación totalmente diferente. Esta variación surge normalmente a raíz de una situación que desestabiliza emocionalmente al docente, por ejemplo, un grupo de alumnos que no se implica, haciéndolo actuar como una "persona diferente".

La teoría dialógica postulada por Hermans (2001, 2009, 2014) explica este hecho como la convergencia entre las dos dimensiones que hemos mencionado, la subjetiva y la social. Del diálogo interno entre estas representaciones y las voces que las configuran, y de forma situacional, emergen lo que se denomina *I-posición* o posiciones-del-yo (Badia y Monereo, s.f.). Estas posiciones-del-yo emergen en la discusión de las voces propias y externas que el sujeto activa para resolver una situación.

De esta manera, desde la identidad del docente, que contiene un conjunto de representaciones que le hacen implementar la evaluación de una manera determinada y relativamente estable en el tiempo, el profesor podría activar posiciones sobre evaluación diferentes según las experiencias en las que se implique, seleccionando entre todas sus representaciones las que crea más adecuadas. Seguidamente, vamos a estudiar cuáles pueden ser las representaciones que pueden constituir los posicionamientos docentes sobre evaluación.

1.5.2. El posicionamiento docente sobre evaluación.

Desde el marco que hemos presentado sobre cómo se entiende la identidad docente, no hemos encontrado estudios en los que se profundice de manera

explícita en el posicionamiento docente sobre evaluación. En este sentido, consideramos que es necesario realizar una primera conceptualización teórica sobre estas posiciones evaluativas. Para ello, vamos a revisar algunos estudios en los que se analizan las concepciones que los docentes tienen sobre la evaluación con el fin de describir cuáles pueden ser sus representaciones y a qué aspectos de la evaluación hacen referencia, aproximándonos de este modo a sus posiciones emergentes sobre la evaluación.

Cuando hablamos de concepciones, podemos definir las como el conjunto de representaciones o estructuras mentales que abarcan creencias, significados, preferencias, etc. sobre un aspecto en concreto. Dice Rodríguez (2010), con respecto a las concepciones sobre evaluación del profesorado:

En la exploración de las concepciones docentes se entiende que la interpretación de sus acciones es producto de la interacción entre los conocimientos sobre las disciplinas escolares, el conocimiento pedagógico y las experiencias en el aula, lo que conforma su conocimiento personal sobre el objeto de reflexión propuesto. Por consiguiente, las ideas, prejuicios y valoraciones de los docentes sobre la evaluación constituyen el marco de interpretación de sus prácticas y permite el reconocimiento del sentido otorgado a través de las respuestas a los interrogantes formulados. (p. 16)

Así pues, examinar cuáles son las concepciones de los docentes sobre evaluación nos permitirá identificar los elementos que conforman su posición evaluativa. El análisis que presentamos en los párrafos siguientes y la Tabla 19, tiene como objetivo realizar una exploración inicial, centrándonos en identificar las categorías en las que se agrupan las representaciones sobre evaluación y el uso que el estudio de estas se da para mejorar la evaluación educativa.

Pero antes, debemos apuntar una consideración: gran parte de la literatura existente bajo la palabra clave concepciones sobre evaluación -*conceptions of assessment*- hace referencia a cuáles son las funciones de la evaluación y cuáles tienen mayor predominancia en el discurso docente. Es decir, investigan las concepciones sobre las funciones de la evaluación, y no el resto de los componentes como el agente de evaluación, el momento de la evaluación, etc.

presuponiendo que el resto de las características se alinean en consonancia con estas funciones. El nombre del cuestionario, *Teachers' Conceptions of Assessment-III* (TCA-III), diseñado por Brown (2004, 2006), en el que se basan la mayor parte de los estudios que adoptan este enfoque, y del cual hemos hablado en el primer apartado cuando describíamos las funciones de la evaluación, pensamos que induce a esta interpretación. Si bien es verdad que las funciones de la evaluación pueden determinar cómo ésta se lleva a cabo, la realidad es que pensar que la evaluación tiene un determinado propósito, no se traduce en una práctica real que dé cuenta de este, como apuntan algunos de los estudios que expondremos a continuación.

Como únicamente estudian las concepciones que tienen los docentes sobre el propósito de la evaluación, no hemos incluido estos estudios en nuestro análisis, aunque sí nos parece importante los resultados que aportan, especialmente para conocer qué utilidad, sentido y propósito otorgan a la evaluación los profesores, cómo ha evolucionado la visión sobre evaluación a lo largo del tiempo y de qué manera influye el contexto en esta y cómo mejorar las formaciones sobre evaluación gracias a los resultados (Barnes et al., 2017; Levy-Vered y Nasser-Abu, 2018; Lumsden y Michaelides, 2011; Muñoz, Palacio y Escobar, 2011; Segers y Tillema, 2011)

Volviendo a los estudios que consideran las concepciones sobre evaluación de los docentes más allá de identificar únicamente la función de la evaluación predominante, a continuación, vamos a exponer, primeramente, un resumen de las investigaciones señalando algunos aspectos como los objetivos del estudio, el diseño metodológico y/o algunos resultados a destacar. Seguidamente, en la Tabla 19 señalamos son los componentes de la evaluación que se investigan en cada estudio, siendo la Tabla 19 un resumen del análisis que se encuentra en el Anexo 4, donde se explicitan las variables que forman parte de cada componente.

En primer lugar, las investigaciones analizadas (Abdul, Venville y Chapman, 2009; Alfageme, Miralles y Torquemada, 2010; Blanco y González, 2010; Dolores y García-García, 2017; Maclellan, 2004; Rueda y Torquemada, 2008; Rodríguez, 2010; Vergara, 2011) nos muestran una variedad de planteamientos y objetivos.

Por un lado, los estudios se centran en describir las concepciones de los docentes sobre evaluación con el propósito de diseñar y/o validar una formación sobre evaluación. La mayoría son estudios de caso en los que se utilizan instrumentos como cuestionarios y entrevistas que se aplican antes y después, y a veces también durante la implementación de las prácticas de evaluación o de la formación sobre evaluación.

Así pues, el conocimiento sobre las concepciones nos ayuda a diseñar y reformular las formaciones sobre evaluación para que realmente respondan a las necesidades de los profesores. El estudio de Alfageme, Miralles y Torquemada (2010), por ejemplo, nos presenta un cuestionario validado sobre las concepciones de evaluación de los docentes de educación secundaria en el área de las ciencias sociales, enfocado a conocer qué, cómo y cuándo evalúa el profesorado; este cuestionario se diseñó y validó con el objetivo de diseñar una formación que aporte mejoras en los procesos de evaluación de los profesores. También es la situación desarrollada en la investigación de Dolores y García-García (2017), en la que, mediante un estudio de casos, se comparan las concepciones sobre evaluación con el enfoque evaluativo propuesto en la reforma, encontrando, por cierto, una notable discrepancia y justificando de esta manera la necesidad de implementar un proyecto formativo que capacite a los docentes a aplicar la nueva reforma educativa de carácter estatal. Por su parte, el estudio de Blanco y González (2010), examina la coherencia entre la evaluación que declaran realizara los profesores y el concepto que tienen sobre esta, llegando a la conclusión de la necesidad de aumentar la formación al respecto.

Además, la descripción de las concepciones sobre evaluación antes y después de las formaciones, nos permite evaluar la eficacia de éstas. Así se realizó en el estudio exploratorio de las concepciones sobre evaluación de Rodríguez (2010), cuando analiza las representaciones antes y después de la formación, encontrando que los docentes están en camino de iniciar cambios para cada una de las categorías, aunque siguen manteniendo una visión tradicional de la evaluación.

Otros estudios, además focalizan su atención en contrastar estas concepciones evaluativas, diferenciando entre aquello que piensan los profesores y aquello que

dicen que hacen. Una manera es construir instrumentos que incluyan preguntas para diferenciar entre aquello que hacen y aquello que piensan que deberían hacer, dentro del mismo plano de la identidad narrada. Por ejemplo, Rueda y Torquemada (2008), en el marco de un taller de formación a profesores universitarios, aplicaron un cuestionario sobre concepciones para recoger la voz de los docentes con preguntas que hacían alusión a cómo piensan que generalmente es la evaluación, lo que piensan que debería ser y lo que piensan que es lo más adecuado. También el estudio de Maclellan (2004) se dirige a preguntar a los docentes cómo creen que debería ser la evaluación, en lugar de examinar cómo la están aplicando.

Por último, encontramos estudios que van más allá y contrastan las concepciones sobre evaluación analizando cómo evalúan en la práctica. La investigación de Vergara (2011), un estudio de casos sobre las concepciones de evaluación de los docentes, integra también en su metodología una observación no participante de situaciones evaluativas y el análisis de documentos, específicamente, de los instrumentos de evaluación. Destacable es que la muestra esté conformada por los docentes que según un sistema de evaluación nacional habían sido considerados como los mejores, con el objetivo de averiguar si sus concepciones están en sintonía con la reforma educativa del momento. El estudio de Abdul, Venville y Chapman (2009) también recogen datos tanto de entrevistas, observación y análisis de documentos en un estudio de casos. Ambos estudios, muestran en sus resultados una disonancia entre lo que creen los docentes que es evaluar y sus prácticas evaluativas.

En segundo lugar, decíamos que la lectura de estos estudios nos permitía conocer cuáles son los elementos que constituyen las concepciones sobre evaluación. En la Tabla 19, podemos leer el análisis realizado: en la primera fila encontramos el total de categorías comunes a todas las investigaciones y, en cada fila, los estudios examinados en orden cronológico y destacando para cada uno cómo se concretaban las categorías -hasta las variables, en el caso que las hubiera-, junto con la referencia de este y los datos principales -nivel educativo y área/s-. Únicamente hemos analizado la información que hace referencia a las concepciones de los docentes sobre evaluación, descartando aquellos aspectos específicos que únicamente eran válidos para una disciplina.

Tabla 19

Representaciones que conforman las concepciones sobre evaluación clasificadas según el componente al que hagan referencia

Investigaciones	Funciones y propósitos	Objeto	Agentes	Proceso (momento) y/o sistema	Instrumentos	Resultados	Cultura de centro y dificultades
Maciellan (2004) Universidad. Varias disciplinas.	Se analiza.	Se analiza.	Se analiza.	Se analiza.		Se analiza.	
Rueda y Torquemada (2008) Universidad. Varias disciplinas.	Se analiza.	Se analiza.		Se analiza.	Se analiza.		Se analiza.
Abdul et al. (2009) Universidad. Matemáticas.	Se analiza.				Se analiza.		Se analiza.
Blanco y González (2010) Educación Primaria.	Se analiza.	Se analiza.		Se analiza.	Se analiza.	Se analiza.	
Alfageme et al. (2010) Secundaria Obligatoria. Ciencias Sociales.	Se analiza.	Se analiza.		Se analiza.	Se analiza.	Se analiza.	
Rodríguez (2010) Educación Preescolar, Básica y Media. Comunicación lingüística.	Se analiza.		Se analiza.		Se analiza.	Se analiza.	
Vergara (2011) Educación básica. Varias disciplinas.	Se analiza.	Se analiza.	Se analiza.	Se analiza.		Se analiza.	Se analiza.
Dolores y García-García (2017) Educación Secundaria. Matemáticas.	Se analiza.	Se analiza.	Se analiza.		Se analiza.		Se analiza.

Nota. Las celdas en gris significan que estos aspectos no se mencionan. Para conocer qué variables se analizan en cada momento se puede consultar el Anexo 4.

Por un lado, uno de los primeros resultados que observamos del análisis de la Tabla 19, es que el propósito de la evaluación siempre es un aspecto analizado. Pensamos que este hecho está relacionado con la creencia de que tener claro por qué evaluamos es uno de los aspectos más representativos de cómo puede ser la evaluación, aunque no por ello el resto de los componentes suelen estar en sintonía. También lo relacionamos con la transición que históricamente ha estado

sufriendo la evaluación, de una función sumativa a una formativa, y el surgimiento de formaciones e investigaciones enfocadas a conseguir este cambio.

Por otro lado, cabe destacar que el número de investigaciones sobre concepciones sobre evaluación -descartando además aquellas que solo se enfocaban a la función como apuntamos más arriba- no es muy elevado. Estamos de acuerdo con Opre (2015) cuando interpela a los investigadores a prestar más atención a la relación entre las características de la evaluación, las concepciones que tienen los docentes sobre esta y las prácticas reales que llevan a cabo en el aula, es decir, a realizar más investigaciones que reconozcan la importancia de las concepciones sobre evaluación para ampliar la comprensión de la evaluación educativa.

En otro orden de cosas, con relación a las temáticas identificadas, las investigaciones contemplan la mayoría de los componentes sobre la evaluación, incluso en aquellos que las respuestas eran más abiertas. Encontramos que en varios casos se sobreponen o que algunas se interpretan de forma más amplia; por ejemplo, cuesta distinguir entre sistema o proceso de evaluación e instrumentos utilizados para evaluar. Aun así, podemos agrupar todas las representaciones descritas en los estudios de la siguiente manera: el propósito de la evaluación, el objeto de la evaluación, los actores de la evaluación, el sistema de evaluación y sus instrumentos, la comunicación de los resultados de la evaluación y la cultura de centro entorno a la evaluación.

Esta clasificación coincide, por una parte, con el marco teórico desarrollado en el primer apartado en referencia a los componentes de la evaluación y la construcción de un sistema evaluativo y, por otra parte, conecta con los componentes de la identidad docente, en tanto que hace referencia a aspectos teóricos -en este caso sobre los propósitos y la definición de la evaluación-, a aspectos más prácticos -de qué manera programar la evaluación para implementarla- y a las atribuciones afectivas -resultados y dificultades asociadas a la evaluación-. Como podemos observar en la Tabla 20, en resumen, dentro del conjunto de representaciones sobre evaluación que nos muestra la literatura analizada, sobresalen representaciones que hacen referencia a las dimensiones que conforman la identidad docente.

Tabla 20

Relación entre las dimensiones de la identidad docente y los componentes de la evaluación que pueden asociarse a cada una de ellas

Dimensiones del posicionamiento en evaluación	Categorías que conforman cada dimensión
Concepciones sobre evaluación	Finalidad de la evaluación Objeto de la evaluación Sistema de evaluación
Estrategias sobre evaluación	Agentes de la evaluación Momento de la evaluación Instrumentos de evaluación
Sentimientos sobre evaluación	Satisfacción respecto a la evaluación Motivaciones sobre evaluación Dificultades en evaluación

La selección de estas categorías coincide también con el marco teórico referente a la evaluación de los aprendizajes desarrollado en los apartados anteriores. Así pues, el posicionamiento de los docentes sobre evaluación está constituido por diversas representaciones sobre el significado de la evaluación, sobre cómo llevarla a la práctica y sobre qué afectos o vivencias le atribuyen. Esta conceptualización nos ayuda a identificar qué posiciones pueden aparecer, profundizando en cómo estas representaciones se interrelacionan y si se constituyen posiciones que podríamos denominar prototípicas en los docentes.

En este sentido, el estudio de Coombs, DeLuca y LaPointe-McEwan (2018), aunque no tiene como propósito analizar la identidad docente, sí refuerza esta idea de identificar cuáles son los perfiles de evaluación según predominen unas variables u otras. Partiendo de las temáticas identificadas como claves en la alfabetización docente en evaluación -propósitos de la evaluación, procesos de evaluación, equidad y teoría de la medición -, los autores diseñaron y aplicaron el inventario denominado *Approaches to Classroom Assessment Inventory (ACAI)* a una muestra de 727 profesores, con el objetivo de analizar los enfoques en evaluación de maestros en formación inicial, en maestros noveles y en maestros con más de cinco años de experiencia, para poder comparar los enfoques según y a lo largo de las etapas profesionales.

Más allá de los resultados, que nos indican que los docentes con menos experiencia priorizan los propósitos sobre la evaluación y la equidad en la misma, rescatamos esta investigación a nivel metodológico, especialmente por su propuesta de utilizar los aspectos claves en la alfabetización sobre evaluación para definir los enfoques en los que declaran situarse los profesores. También porque en el diseño del instrumento, los autores proponen diferentes escenarios en los que se plantean situaciones, basadas en dilemas reales sobre evaluación, que los maestros deben responder escogiendo entre varias opciones para cada una de las temáticas descritas como claves en la alfabetización docente en evaluación. Es decir, los maestros deben posicionarse y explicitar qué concepciones tendrían en cuenta para resolver un dilema, que puede resultar un incidente crítico en la realidad, sobre evaluación.

Con todo, después de identificar qué categorías constituyen las diferentes dimensiones del posicionamiento docente sobre evaluación, es necesario concretar qué variables pueden formar parte de cada una. Por ejemplo, en relación con la función de la evaluación, como hemos señalado en el análisis de la literatura, los docentes pueden posicionarse en una evaluación centrada en acreditar, en rendir cuentas, en mejorar la enseñanza o en ayudar a los alumnos a aprender. Según las representaciones que declaren, construirán una posición u otra.

En la Tabla 21 detallamos las representaciones que pueden constituir cada una de las categorías claves del posicionamiento docente sobre evaluación. Recordemos que las categorías se agrupan según las dimensiones de la posición en evaluación, de la siguiente manera: la finalidad, el objeto y el sistema se corresponden con la dimensión de concepciones sobre la evaluación; los agentes, el momento y los instrumentos se corresponden con la dimensión de estrategias sobre la evaluación; y la satisfacción, las motivaciones y las dificultades se corresponden con la dimensión de sentimientos sobre la evaluación.

Además, el conjunto de variables puede leerse clasificadas en dos bloques: uno que reúne la mayoría de las variables que son representativas de una evaluación de carácter tradicional, y otro que reúne aquellas que son más coincidentes con una evaluación de carácter auténtico. El hecho de identificar estos dos enfoques

responde a la idea desarrollada en el marco teórico sobre los tipos de evaluación, donde demostramos que en la literatura existente se establece una dicotomía entre evaluación tradicional y evaluación alternativa, dándose claramente una evolución hacia la segunda. Y, de la evaluación alternativa, concretamos en evaluación auténtica por ser el modelo en el que se basa esta investigación, justificado en el apartado sobre evaluación auténtica.

Tabla 21

Continuo de variables que definen las categorías del posicionamiento docente sobre evaluación

Categorías	Variables que definen un enfoque de evaluación tradicional		Variables que definen un enfoque de evaluación auténtico	
Finalidad	Función de acreditación	Función de rendición de cuentas	Función de mejora de la enseñanza	Función de mejora de los aprendizajes
Objeto	Conocimientos según valoración normativa	Conocimientos según valoración criterial	Competencias y proceso de enseñanza	Competencias y estrategias de aprendizaje
Sistema	Programación diseñada por una editorial u otro servicio	Programación según las exigencias institucionales	Sistema diseñado por el docente	Sistema consensuado a nivel de centro
Agentes	El docente cumpliendo indicaciones de agentes externos	El docente cumpliendo indicaciones de agentes internos	El docente según su criterio y contexto de aula	El docente y los alumnos participan en la evaluación
Momento	Al finalizar la programación	Al finalizar la programación	Durante y después	Antes, durante y después
Instrumentos	Exámenes de respuesta única	Exámenes basados en ejercicios y problemas	Resolución activa de problemas auténticos	Tareas auténticas que permitan reflexionar sobre su resolución
Satisfacción	Niveles siempre fijos y/o indiferencia	Niveles siempre fijos y/o indiferencia	Niveles que varían según el proceso	Niveles que varían según el proceso
Motivación	Impartir justicia	Demostrar el trabajo realizado	Mejorar la enseñanza	Mejorar el aprendizaje de los alumnos
Dificultades	Adjudicar una calificación justa	Valorar los resultados como indicador	Recoger información para ajustar la ayuda educativa	Aportar evidencias e instrumentos de autorregulación

Cabe señalar que estas variables son acumulativas entre sí, de forma que tender hacia una evaluación auténtica no significa abandonar el resto de las variables, sino

apostar por priorizar una evaluación orientada al aprendizaje y al desarrollo de competencias en una escuela inclusiva, como hemos venido defendiendo a lo largo de este estudio. Es decir, posicionarse, por ejemplo, en una evaluación de las estrategias de los alumnos para ayudarles en su autorregulación no significa dejar de evaluar conocimientos, ni siquiera dejar de acreditarlos a través de una evaluación final que permita medir el nivel alcanzado. Tanto la diferenciación en dos bloques como la continuidad de los mismos, coincide con la propuesta de otros estudios, como por ejemplo el de Postareff, Virtanen, Katajavuori, Lidblom-Yläne (2012), en el que las categorías sobre las concepciones forman un continuo que va desde las concepciones más reproductivas, con descripciones y reflexiones bastante limitadas y relacionadas con una principalmente visión acreditativa, a concepciones transformadoras, con descripciones más profundas y reflexivas que incluyen una visión más pedagógica.

Además, debemos considerar esta descripción como un marco para explorar si existen posiciones evaluativas prototípicas, para lo cual sería necesario diseñar una investigación que nos permita identificarlas y describirlas, como realizaremos más adelante. Avanzándonos en esta hipótesis, es de esperar que haya una coherencia entre las variables. Por ejemplo, si los docentes declaran utilizar instrumentos de carácter formativo como el portafolio, también es de esperar que declaren que en la evaluación deben participar también los alumnos.

Bajo este supuesto, podríamos encontrar diferentes posicionamientos según las categorías con mayor predominancia. Por ejemplo, podríamos definir la posición evaluativa centrada en medir los aprendizajes, descrita de la siguiente manera: partiendo del propósito de acreditar -función-, centrada en los conocimientos -objeto-, diseñada únicamente por el profesor -sistema-, el cuál es el encargado de aplicarla -agente-, al finalizar la unidad didáctica -momento- mediante un test de respuesta múltiple -instrumento-, sin ganas de repensar este proceso -motivación- y descubriendo que los alumnos no son capaces de regularse -dificultades-, resultado que piensa que se debe a lo poco que se esfuerza el alumnado -atribución-. Obviamente, esta descripción que hasta hace poco era una realidad constante en las aulas, es posible que nos parezca ahora hasta casi caricaturesca.

Porque, si bien es verdad que las concepciones sobre evaluación influyen en la práctica cierto es también que existe una brecha importante entre ambas, de forma que las prácticas de evaluación pueden no ser consecuentes con sus concepciones (Hidalgo y Murillo, 2017). Es decir, debemos considerar también la posibilidad de que la posición evaluativa no sea tan clara en su coherencia entre variables y encontremos multiplicidad de posiciones según sea la realidad de la práctica evaluativa de cada docente.

Es interesante pues, extender la investigación y realizar estudios que analicen de forma más profunda cómo es este posicionamiento puesto en práctica, para descubrir cuáles son las voces con las que dialogan los docentes en su posición en acción respecto a la evaluación. Un análisis desde una teoría dialógica nos permitiría conocer al detalle estos posicionamientos, con el fin de ayudar al docente a superar ciertos incidentes y a acompañarlos en el desarrollo de su competencia en evaluación. Por ello, en el siguiente apartado vamos a destacar cuáles pueden ser los dispositivos de cambio que hagan que el docente transite entre varias representaciones y reconstruya su posicionamiento sobre la evaluación, en la dirección de una evaluación auténtica.

1.5.3. Dispositivos para la reconstrucción del posicionamiento en evaluación

Definíamos la competencia en evaluación de los profesores como el conjunto de conocimientos y habilidades sobre evaluación y su capacidad estratégica para utilizarlos adecuadamente en su contexto, hecho que determina la eficacia en sus prácticas evaluativas. Dicha competencia queda manifiesta en aquello que piensan que es la evaluación y en cómo deberían aplicarla, así como en su respuesta emocional asociada. También se refleja durante su actividad evaluadora cuando, condicionada por el surgimiento de dificultades o problemas, consiguen o no ajustarse a la situación. Es decir, analizando la posición sobre evaluación, declarada o en acción, podríamos describir cuál es la competencia en evaluación de los docentes, como punto de partida para trabajar por mejorarla.

Además, podemos afirmar que cuando el conjunto de representaciones que sostienen sobre la evaluación es coherente entre ellas y conforman un posicionamiento enfocado al marco actual de enseñanza competencial, es decir, con una clara tendencia a la evaluación auténtica, el profesorado obtendrá resultados positivos en su tarea de evaluar. Sin embargo, no siempre sucede así y, por un lado, las investigaciones nos muestran que existen discrepancias entre lo que se piensa y lo que se hace en evaluación (Abdul et al., 2009; Vergara, 2011) y, por otro lado, poca coherencia entre las propias representaciones debido al desconocimiento de los componentes de la evaluación y sus características.

Por tanto, se hace evidente la necesidad de acompañar a los docentes a sostener un posicionamiento sobre evaluación coherente, que les permita implementar una práctica evaluativa realmente auténtica de forma estratégica. Para ello, optamos por realizar una intervención que impacte en este posicionamiento, haciendo que el docente tome consciencia de cuál es y sobre qué dimensiones de su identidad debe trabajar para redefinirlo con el fin de ajustarse a los contextos diversos y cambiantes en los que tendrá que evaluar. Queda justificada la idea de que, si queremos conseguir que los docentes sean competentes en evaluación auténtica, no es suficiente con dotarles de las estrategias evaluativas -aunque pueda ser un primer paso-, sino que es necesario, además, incidir en su identidad docente, de forma que cuestionen su posicionamiento sobre la evaluación con el fin de poder reconstruirlo. Hemos identificado dos dispositivos que nos permiten este objetivo, el primero consiste en ofrecer instrumentos para que puedan cuestionar y transformar sus prácticas evaluativas, y el segundo, a partir de este cambio y de los incidentes que puedan surgir, tomar conciencia su posicionamiento sobre evaluación y cómo reconstruirlo. A continuación, describiremos cada dispositivo y lo justificaremos basándonos en la literatura al respecto.

En primer lugar, uno de los dispositivos de cambio consiste en implementar un proyecto formativo que impacte en la manera de pensar, hacer y sentir la evaluación del docente. Como señalan Liesa, Castelló y Becerril (2018), para que los docentes puedan cambiar sus prácticas es necesario que reconstruyan sus concepciones y para ello, deben poder nombrarlas, comprenderlas y tomar consciencia y, por tanto, deben aprender sobre, en nuestro caso, la evaluación. De esta forma estarán más

predispuestos para analizar sus prácticas y, por tanto, transformarlas (Carretero, Liesa, Mayoral y Mollà, 2008). Además, existen experiencias previas en las que, después de una formación sobre evaluación, los docentes se apropiaron de nuevas voces pedagógicas sobre evaluación, transformando su discurso y enfocándolo a una evaluación orientada al aprendizaje (Badia y Chumpitaz-Campos, 2018). También, como ya hemos desarrollado en anteriores apartados, existen experiencias que demuestran que después de una formación los docentes cambian sus prácticas y/o sus concepciones sobre evaluación (Monereo, Castelló et al., 2009; Rodríguez, 2010).

Por lo tanto, hay evidencias que demuestran que, si queremos conseguir que los cambios en evaluación que realicen los profesores sean sostenibles en el tiempo, una de las primeras acciones es implementar una formación; es decir, el primer paso es que sepan cómo realizarlos. Además, para provocar un impacto, creemos que es necesario que los docentes deban poner en práctica una nueva manera de evaluar, enfrentándose a los desafíos que esta nueva situación puede generarles. El Proyecto GAPPISA, centrado en la evaluación auténtica, con su guía y su ciclo formativo basado en la acción-reflexión, se presenta como una muy buena propuesta para cumplir este objetivo. Además, la formación nos permite recoger las concepciones de los docentes antes y después, la creación de un instrumento de evaluación, y su aplicación favorece el análisis de las estrategias y sentimientos asociados. Por último, el reto de aplicar algo nuevo puede ocasionar momentos desestabilizantes para el posicionamiento docente sobre evaluación, hecho que nos posibilita poder proponer un nuevo dispositivo de cambio: el análisis de los incidentes críticos sobre evaluación como propuesta formativa.

En segundo lugar, partiendo de las posibles problemáticas y dificultades derivadas de implementar una práctica innovadora como es la evaluación auténtica y reconociéndolas como una oportunidad de aprendizaje, el análisis de ésta se presenta como el siguiente paso para conseguir que los profesores reconstruyan su posicionamiento sobre evaluación. El análisis de los incidentes críticos vividos por los profesores es una buena estrategia en este sentido (Monereo, Weise y Álvarez, 2013; Monereo y Monte, 2011; Weise y Sánchez-Busqués, 2013). Entendemos por incidente crítico una situación acotada en el tiempo y en el espacio

que desestabiliza emocionalmente al profesor, de forma que pone en crisis su posición identitaria (Monereo, Badia et al., 2009). Este suceso desestabilizante puede tener como resultado tanto que el profesor se reafirme en las concepciones, estrategias o sentimientos que le impiden resolverlo, como que, a partir de ese momento, decida trabajar en un cambio que le permita revisar algunos aspectos de su identidad profesional.

El análisis de los incidentes críticos ha resultado ser una muy buena estrategia de aprendizaje para el profesorado. En este sentido, uno de los instrumentos más destacados es la Pauta para el Análisis de Incidentes Críticos (PANIC) de Monereo (2010) a través de la cual los docentes describen su incidente y, junto a un asesor, trabajan en su intervención y seguimiento. Son bastantes los estudios que demuestran que, en la formación inicial del profesorado, analizar los incidentes críticos que son comunes a la docencia y trabajar sobre cómo afrontarlos, permite a los estudiantes universitarios desarrollar sus competencias profesionales (Del Maestro y Monereo, 2014; Fernández y Marrero, 2013).

Además, se ha demostrado su eficacia en profesores en ejercicio a través de la interpretación, a posteriori, de los incidentes críticos que han experimentado. En el contexto universitario, encontramos el estudio de Weise y Sánchez-Busqués (2013) en el que se analiza la interpretación que los profesores hacen de incidentes críticos que han vivido, dando como resultado la toma de conciencia del uso de estrategias de resolución poco eficientes por su parte y, por tanto, la necesidad de continuar formándose. También la formación sobre incidentes críticos evaluada en Monereo, Weise y Álvarez (2013), en la que la metodología consistía en dramatizar los incidentes para aprender a afrontarlos y que tuvo por resultado cambios en algunas dimensiones de la identidad docente, de manera sostenible en el tiempo.

En el contexto de formación de maestros de educación primaria en ejercicio también se ha demostrado la eficacia del uso de incidentes críticos como estrategia de formación continua (Bilbao y Monereo, 2011; Nail et al., 2012). De estos estudios se deduce que, en lugar de formar sobre una temática en concreta, una manera eficaz de que los resultados de la formación permanente sean positivos, es orientar la formación de maestros hacia el estudio de los incidentes críticos, aprovechando

su análisis consciente para aprender estrategias efectivas, en nuestro caso, sobre evaluación.

Relacionando ahora, incidentes críticos y evaluación, encontramos que dentro de los estudios que describen cuáles son los incidentes más comunes, los relativos a la evaluación conforman un grupo propio de incidentes. En algunos casos, se aprende a partir de ellos, como en el estudio de Elórtégui, Medina y Fernández (2003) en la que se examina la siguiente situación: un estudiante nada problemático en el aula y que durante la enseñanza se implica activamente, el día del examen, pasados diez minutos después de su recepción, lo entrega en blanco. A partir de este planteamiento, los participantes discuten sobre diferentes alternativas generando muchas cuestionen que pueden poner en duda el sistema de evaluación y posibilitar un cambio en este sentido.

Por último, encontramos en el artículo de Monereo (2019), una primera aproximación sobre cuáles son los principales incidentes que afrontan los centros educativos con relación a la evaluación de los aprendizajes y de qué maneras pueden enfrentarse. Los incidentes se clasifican en cinco grandes bloques:

- 1) Dificultades para compartir los criterios de evaluación con los compañeros. En este caso el profesorado da muestras de poca coordinación y consenso en el momento de acordar unos principios mínimos que presidan sus decisiones. Por ejemplo, falta de equidad, de objetividad, de consenso o de personalización.
- 2) Dificultades para superar conflictos éticos. Aquí se agruparían decisiones y conductas que atentan contra la actuación docente competente. Por ejemplo, inmovilismo de algunos docentes, necesidad de formación, imposiciones que bloquean o comentarios poco éticos hacia los alumnos.
- 3) Dificultad para diferenciar entre calificación y fracaso. Formado por aquellas situaciones, herederas de una evaluación tradicional, en las que la calificación es la única estrategia para dar cuenta del aprendizaje, y además se utiliza para juzga, clasificar y presionar. Por ejemplo, competición y

negociación por la nota más alta, familias que protestan o alumnos que se sienten tan presionados por obtener una calificación que acaban mintiendo.

- 4) Dificultades para evaluar competencias. Un conjunto formado por situaciones en las que es resulta complejo que metodología y evaluación estén en sintonía. Por ejemplo, a causa de la fragmentación del currículum, por falta de consenso o de formación.
- 5) Dificultades con ajustar las expectativas. Cuando el profesorado espera mucho más de lo que está obteniendo. Por ejemplo, descubrir deficiencias en el sistema de evaluación o que esta no ha servido para aprender.

Las resoluciones a estos incidentes críticos sobre evaluación propuestas por el mismo Monereo (2019), se basan en la implementación de planes, programas y proyectos, como, por ejemplo, plan de creación de bancos de recursos, diseño de proyectos sobre mediación de conflictos y creación de planes que establezcan protocolos de actuación. Estas medidas están enfocadas a construir una cultura evaluativa de centro que permita que las innovaciones introducidas en evaluación sean permanentes y estables en el tiempo, gracias a la creación de una comunidad que respalda y promueve dicho cambio.

En resumen, el posicionamiento sobre la evaluación que tienen los profesores ha de entenderse en la interrelación entre rol docente sobre evaluación, las estrategias evaluativas y las atribuciones emocionales asociadas. Y, provocar que los docentes introduzcan prácticas de la evaluación auténtica, posibilita hacer emerger las representaciones sobre evaluación, a la vez que incidentes críticos que son una buena oportunidad para reconstruir su posicionamiento, siendo tanto la formación sobre evaluación como el estudio del posicionamiento y del análisis de los incidentes críticos sobre evaluación, una buena manera de formar docentes competentes en evaluación.

Capítulo 2. Estudio 1: Evaluando las Evaluaciones. Eficacia de una Guía para la Elaboración de Pruebas Auténticas

2.1. Introducción y Objetivos

La evaluación tiene una influencia innegable sobre los procesos de enseñanza y aprendizaje. El hecho de determinar qué queremos valorar y con qué propósito lo hacemos, condiciona desde la programación de actividades hasta las estrategias de aprendizaje de los alumnos. Cuestionar la evaluación, según la ya popular máxima “dime cómo evalúas y te diré cómo aprenden tus alumnos”, puede suponer una estrategia muy potente para avanzar en un enfoque por competencias dados los efectos retroactivos que puede tener el hecho de modificar la evaluación sobre los procesos de enseñanza y aprendizaje.

Sin embargo, y a pesar del progresivo cambio de paradigma psicoeducativo hacia una educación por competencias (Comisión Europea, 2007), continúa dándose una gran resistencia para cambiar los métodos de evaluación (Álvarez, 2001). Esto sucede al menos por dos motivos: en primer lugar, la gran distancia existente entre una evaluación conceptual, memorística y verbal, centrada en recoger qué sabe decir el alumno, y una evaluación estratégica, comprensiva y competencial, centrada en promover la resolución de problemas significativos (Santos-Guerra, 2003); y, en segundo lugar, una intencionada separación entre enseñar y evaluar, bajo el convencimiento de que la evaluación más “objetiva” es la que se produce de forma descontextualizada.

Para avanzar hacia el modelo de evaluación de competencias, primeramente, debemos compartir el significado de competencia. Según De la Orden (2011) una competencia es el conjunto de conocimientos y estrategias que permiten a una persona, en calidad de ciudadano, profesional, consumidor, etc., afrontar problemas relevantes de ese ámbito de desarrollo, es decir, el desempeño de un rol de forma exitosa. A nivel de diseño de actividades de aprendizaje y de evaluación esto supone definir los contextos y prácticas en las que los aprendices participarán como futuros ciudadano y las condiciones, elementos y tareas en las que se verán implicados.

Aparejado a este modelo de desarrollo de competencias, se hace imperioso utilizar unas prácticas de evaluación de carácter dinámico, capaces de acompañar en lugar de solo juzgar dicho desarrollo. Es decir, una evaluación que se concrete en unas actividades conectadas con la realidad y que permitan reflexionar sobre las respuestas aplicadas. La propuesta que de forma más ajustada y coherente responde a esta perspectiva es la denominada evaluación auténtica, que plantea valorar las competencias “en activo”, es decir, a través de su ejecución en la resolución de problemas representativos y fidedignos de la realidad.

En la abundante literatura sobre evaluación auténtica publicada (Ashford-Rowea et al., 2014; Brown, 2015; Darling-Hammond y Snyder, 2000; Gulikers et al., 2006; Frey et al., 2012; A. Herrington y J. Herrington, 2006; Villarroel et al., 2017) se describen las tres categorías que la definen. Primero, la necesidad de hacer referencia a un problema auténtico que adopte la forma de una tarea compleja en la que en su resolución se requiera una mayor o menor disposición de recursos y de exigencia cognitiva, es decir, un mayor o menor grado de competencialidad. Segundo, su grado de inclusividad, en el cual la evaluación admite distintas formas y modalidades de respuesta y variadas formas de demandas, permitiendo que los alumnos puedan abordarla. Tercero, que la evaluación sea capaz de promover aprendizaje, lo que llamamos su poder formativo, por cuanto este tipo de pruebas favorece y promueve la búsqueda de alternativas de respuesta y el aprendizaje a partir de los propios errores. A continuación, pasamos a definir cada una de estas categorías sobre las cuales girará nuestro estudio.

La competencialidad se define por la tipología de contenidos que se tratan y los niveles de exigencia cognitiva en los que se sitúan estos. Hablamos entonces de complejidad cognitiva, entendida como el análisis de las tareas considerando el tipo de habilidades de pensamiento que se demandan para poner en práctica los saberes adquiridos, tomando de referencia los niveles definidos en la taxonomía de Bloom tal y como Tufail, Murtaza y Iqbal (2016) plantean en su estudio. Considerando esto, la selección de conocimientos que se pretenden evaluar debe coincidir con aquellos saberes implicados en la solución del problema y concretarse en tareas acordes al nivel de complejidad que este exige (Darling-Hammond y Snyder, 2000). Por ejemplo, no podemos evaluar la capacidad estratégica de un

aprendiz para publicar un artículo en un blog especializado, si únicamente le pedimos que defina qué es un artículo de divulgación científica. Además, la exigencia cognitiva debe ir aparejada con la misma que se requeriría en la situación real, planteando las tareas de forma que su realización pase por una movilización progresiva de los recursos cognitivos (Avery et al., 2012). Continuado con el ejemplo anterior, seguir todos los pasos implicados en la publicación, desde la investigación, pasando por las estrategias de escritura hasta la revisión del proceso cuando ya se ha publicado. Dicho de otro modo, para que la resolución sea constructiva tiene que existir una secuencialidad e interrelación entre las tareas propuestas (Meyers y Nulty, 2009) de forma que durante el proceso se puedan desarrollar las competencias necesarias.

Por último, esta característica también hace referencia al carácter socializador de la evaluación, puesto que las actividades competenciales implican posicionarse en un rol específico dentro de la comunidad en la que se sitúa la práctica (Guliker et al., 2008), por lo tanto, en el diseño de la evaluación deben seleccionarse las características físico-sociales, la aplicabilidad y los conocimientos previos necesarios que envuelven el problema real, para recrearlas siempre que sea posible.

En cuanto a la inclusividad, esta parte del principio de equidad y por lo tanto de la importancia de ofrecer adaptaciones y mecanismos para que, con independencia de su origen cultural, sus capacidades, etc., la mayor parte del alumnado tenga la oportunidad de acceder y participar en la resolución de la tarea de evaluación (Kaur et al., 2015). Así pues, una evaluación inclusiva es aquella que facilita a todos los aprendices, a partir de sus necesidades (Coll y Onrubia, 2002), diferentes ayudas para demostrar su comprensión y adquisición de competencias. Se trata de ajustar las condiciones de resolución, y no de simplificar la complejidad del problema eliminando o minimizando contenidos. La exigencia cognitiva debe mantenerse, pero el acceso a la información, su elaboración y la comunicación o ejecución de la respuesta, pueden personalizarse.

Para conseguirlo se deben ofrecer ayudas para guiar la resolución, sea mediante, a) variando el soporte en el que se presenta la información (de manera textual,

visual, audiovisual...) o los materiales de apoyo para acceder a ella (plantillas, pautas...); b) compartiendo los sistemas de valoración de las distintas decisiones y acciones (a través de rúbricas de calidad, bases de orientación, consejos... de los procesos ejecutados y los productos obtenidos); o, c) ofreciendo la oportunidad de resolver las situaciones inesperadas de forma colaborativa. Estas medidas, además, posibilitan que la práctica de evaluación se desarrolle en un ambiente que los aprendices valoran como social y emocionalmente positivo y seguro (Kaur et al., 2017).

Respecto a la cualidad o potencialidad formativa del problema o tarea de evaluación, se pone en práctica cuando se considera la profundidad en que promueve la reflexión sobre la ejecución efectuada, generalmente a través de pautas para hacerlo. Con ello, no solo se evalúan las competencias activadas por el aprendiz, sino que también se le prepara para prácticas de desempeño futuras (Raymond et al., 2013) y se favorece la toma de conciencia de todo el aprendizaje logrado. Aunque los instrumentos de evaluación auténtica pueden ser múltiples, los que dan más cuenta de esta característica son los portafolios, con soporte en papel o digital (Sabirón y Arraiz, 2013), las rúbricas dirigidas a destacar los procesos subyacentes de la resolución (Dawson, 2015) y la coevaluación y autoevaluación auténticas (Kearney, Perkins y Kennedy-Clark, 2016).

Si bien consideramos que la necesidad e interés de utilizar modalidades de evaluación auténtica para valorar las competencias de los aprendices está suficientemente argumentada y justificada, un tema mucho menos estudiado y, sin embargo, igualmente crucial, es el referido a las competencias que tienen los docentes para diseñar, desarrollar y revisar pruebas de evaluación auténtica. Estar alfabetizado en evaluación significa saber elaborar, aplicar y reelaborar instrumentos de evaluación adecuados a los objetivos que se formulan, así como conocer el potencial que tiene la evaluación y comprender los efectos que provoca para poder usarlos en pro de la mejora de la enseñanza (Deluca et al., 2016). Esta competencia en evaluación debe ajustarse al modelo actual de evaluación y, por lo tanto, los docentes deben actualizar sus concepciones y ponerlas al servicio de las nuevas prácticas tanto de enseñanza como de evaluación (Sanahuja y Sánchez-Tarazaga, 2018; Tonda y Medina, 2013).

Desgraciadamente, dicha competencia no ha tenido el reconocimiento que se merece en los planes de formación inicial del profesorado (Popham, 2009; Stiggins, 2002) y los modelos de evaluación que se utilizan habitualmente para medir la competencia docente tampoco han considerado la competencia en evaluación entre sus escalas (Goodman et al., 2008). Por fortuna, en la actualidad, empiezan a existir propuestas que implican a los futuros docentes en formación, en calidad de aprendices, en sistemas de auto y coevaluación de sus prácticas profesionales, y en calidad de enseñantes, en la confección de pruebas de evaluación competenciales dirigidas a sus alumnos (Gómez-Devís et al., 2017). En todo caso, el debate sobre la evaluación, en el interior de los equipos docentes, requiere todavía mucha más presencia, en tanto que todavía existe profesorado que declara que les faltan herramientas para poner en acción los nuevos enfoques evaluativos (Ramírez, 2016).

Este escenario ha provocado que exista una demanda creciente de actividades formativas y proyectos de innovación encaminados a propiciar cambios en la evaluación tradicional. Sin embargo, el tipo de formación que suele ofrecerse tiene un carácter de implicación individual y finalmente afecta a pocos profesores, docentes que además suelen ya estar en línea con los nuevos enfoques. Estos mismos profesores, tienen pocas posibilidades de cambiar las resistencias de los respectivos claustros a los que pertenecen, compañeros que mayoritariamente recelan de los nuevos paradigmas evaluativos y que difícilmente abandonan su “zona de confort” (Ahumada, 2005; Syrene et al., 2009). Por otro lado, esas mismas propuestas formativas no suelen actuar en lo que Vygotsky (1978) denominaría “la zona de desarrollo próximo” del profesor. Este espacio se define como aquel que permitiría un avance real y consistente en sus prácticas evaluativas, al partir de lo que el docente ya lleva a cabo cotidianamente para moverlo a una etapa posterior.

Precisamente, con intención de responder a estas carencias, la investigación que planteamos pretende mostrar los beneficios de una propuesta formativa que: a) sea institucional y, por consiguiente, actúe sobre equipos docentes y no solo sobre individuos aislados; y b) parta de las prácticas evaluativas reales que los docentes ya realizan en sus respectivas asignaturas con el fin de lograr pruebas de evaluación más auténticas, en un proceso que precisamente denominamos

autenticación de pruebas de evaluación. La finalidad última es crear redes profesionales inter-centros que apoyen y promuevan la innovación en la evaluación de competencias. En estudios preliminares realizados por nuestro equipo (Monereo, Castelló et al., 2009) hemos obtenido indicios que avalan este planteamiento.

Para finalizar, los objetivos de esta investigación van dirigidos al análisis de la calidad de los instrumentos de evaluación diseñados por los participantes, así como a la validación del propio instrumento formativo. El análisis está enfocado a caracterizar las pruebas de evaluación realizadas y comprobar si son efectivas para evaluar competencias, producir nuevos aprendizajes y ajustarse a los diferentes perfiles de aprendizaje, es decir, si son competenciales, formativas e inclusivas, como hemos explicado anteriormente. Así pues, los objetivos fueron los siguientes:

1. Lograr una mejora significativa en el grado de competencialidad de las pruebas elaboradas por los docentes, tras la formación.
2. Conseguir una mejora significativa en el potencial formativo de las pruebas elaboradas por los docentes, tras la formación.
3. Obtener una mejora significativa en la capacidad de inclusividad de las pruebas elaboradas por los docentes, tras la formación.
4. Analizar el tipo de cambios producidos en cada dimensión y valorar las fortalezas y debilidades del instrumento para autenticar pruebas (la guía GAPPISA).

2.2. Método del Estudio

2.2.1. Contexto.

La presente investigación se enmarca en la formación a docentes sobre evaluación auténtica que denominamos Proyecto GAPPISA (Mejías y Monereo, 2018), coincidiendo con el acrónimo que da nombre al instrumento principal, *Guía para el Análisis de Pruebas desde la perspectiva PISA* (Álvarez-Bernárdez y Monereo, 2016; Duran y Monereo, 2008; Monereo, Castelló et al., 2009; Monereo y Morcillo, 2011). Aunque en el marco teórico ya hemos descrito en profundidad esta formación sobre evaluación auténtica, a continuación, vamos a recordar los aspectos principales.

El objetivo de la formación fue dotar de estrategias y recursos a los docentes para que pudieran modificar y diseñar actividades evaluativas que cumplieran los criterios característicos de la autenticidad, fundamentales para la evaluación de competencias. La formación consistió en el diseño y validación de pruebas de evaluación competenciales, inclusivas y formativas, mediante el análisis del conocimiento teórico, el uso de instrumentos auto-formativos, la aplicación en actividades de evaluación reales y la mediación directa de un asesor especializado en evaluación. La actividad principal consistió en que los docentes eligieran una de las pruebas de evaluación que utilizaran más asiduamente, la modificaran para hacerla más auténtica, la aplicasen en una situación de evaluación real con su grupo de alumnos y reflexionaran sobre los resultados logrados.

Para llevar a cabo esta formación docente, el Proyecto GAPPISA se organizó en cuatro fases: un primer momento en el que los profesores participaron de forma presencial en unas jornadas de formación teórico-práctica sobre la evaluación auténtica y la transformación de pruebas evaluativas; una segunda fase en la que analizaron sus pruebas y las re-elaboraron en base a las características principales definidas en la teoría sobre evaluación auténtica; una tercera fase en la que los profesores aplicaron la prueba diseñada en una situación real; y una última fase donde se reflexionó sobre los cambios producidos a consecuencia de la aplicación y se presentaron los resultados y la evaluación de la formación. La duración del

curso fue anual de forma que las diferentes actividades formativas se distribuyeron a lo largo de los tres semestres escolares con el propósito de que se pudieran realizar las actividades progresivamente.

Durante este proceso, además de los materiales formativos diseñados y puestos a disposición de los participantes mediante una plataforma didáctica con actividades, recursos, ejemplos etc., los asesores pidieron evidencias sobre los productos realizados. De esta manera, los docentes debían enviar la prueba original de partida, la primera versión de la prueba autenticada y la prueba final, aplicada a su grupo de clase. A su vez, el asesor enviaba una retroalimentación individualizada de cada prueba elaborada para que los docentes pudieran mejorarla.

2.2.2. Participantes.

El Proyecto GAPPISA se llevó a cabo durante dos cursos escolares consecutivos en una red de escuelas privadas distribuidas por diferentes comunidades autónomas de España -Andalucía, País Vasco, Zaragoza, Navarra, Castilla y León y La Rioja-. Los participantes se inscribieron de forma voluntaria, con un mínimo de cinco personas por centro, que superamos con creces en muchos casos, así que partían todos de un claro interés por mejorar sus prácticas de evaluación. Contamos con una muestra final de 206 participantes, pertenecientes a 17 centros educativos, con las siguientes características: 60% mujeres y 40% hombres; estaban representadas todas las materias educativas, dominando las lenguas, matemáticas y ciencias sociales -aproximadamente un 70%-; los cursos escolares iban de infantil hasta educación secundaria obligatoria, con una mayor representación de esta última etapa -58,8%-; el 52% de los docentes tenía una experiencia profesional menor a 15 años y el 64% una experiencia mayor a 15 años.

2.2.3. Procedimiento.

Los momentos de recogida de datos coincidieron con las fases de implementación del proyecto de formación. En la primera fase, los docentes enviaban la versión original de la prueba de evaluación sobre la que querían trabajar. Cada profesor elegía la prueba que deseaba modificar en función de dos parámetros: que fuese muy relevante para alcanzar los objetivos de la asignatura y que no fuese

competencial. En la segunda fase, recogimos la primera versión de su prueba original. Esta prueba era la prueba original modificada después de analizarla con la Guía GAPPISA e introducir cambios recomendados en la teoría sobre evaluación auténtica trabajada con anterioridad; a este proceso le llamamos “proceso de autenticación de la prueba” y a esta segunda versión recogida prueba autenticada. En una tercera fase, después de la retroalimentación realizada por el asesor sobre la segunda versión, recopilamos las últimas versiones de la prueba que denominamos prueba fina, que a posteriori deberían aplicar en su grupo clase ese mismo año.

2.2.4. Instrumentos.

Seguidamente pasamos a describir los instrumentos de formación y evaluación utilizados: la Guía GAPPISA y la Rúbrica de evaluación de pruebas auténticas.

Como hemos comentado anteriormente, el instrumento principal es la Guía GAPPISA, la cuál parte de trabajos anteriores donde se desarrolló e implementó durante un asesoramiento a profesores con el objetivo de que reformularan sus pruebas y exámenes tradicionales (Monereo y Lemus, 2009; Monereo, Castelló et al., 2009).

Actualmente está compuesta por diez características que, desarrolladas en mayor o menor grado, determinan la autenticidad de la prueba de evaluación o actividad que diseñe el docente según su competencialidad, potencial formativo e inclusividad. En el capítulo de introducción se expone una tabla (ver Tabla 18) en la que describimos cuáles son las categorías principales, las preguntas que orientan la transformación y una breve descripción de estas. Además, en el Anexo 1 se encuentra la versión de la Guía GAPPISA en papel.

Sobre la rúbrica de evaluación de pruebas auténticas (disponible en el Anexo 2), esta se utilizó para la revisión de pruebas por parte del equipo de asesores con el fin de realizar una retroalimentación a los participantes. Está elaborada en base a las categorías que conforman el modelo de autenticación de pruebas evaluativas, agrupando el alcance de cada característica en cuatro niveles distintos de adquisición según los cambios realizados. De esta manera, con el instrumento

podemos determinar la puntuación de cada característica y del total de la prueba según el nivel de logro alcanzado. Además, incluye una puntuación punitiva en cada una de las dimensiones, la cual indica que el cambio no se está haciendo en la dirección correcta, es decir, que se está malinterpretando la característica e introduciendo cambios que, en lugar de resultar positivos, pueden ser irrelevantes, superficiales o perjudiciales.

2.2.5. Análisis de datos.

Para analizar las pruebas, empleamos la rúbrica de evaluación de pruebas auténticas en la comparación de las dos pruebas de las que disponíamos en cada momento -prueba original con prueba autenticada y prueba autenticada con prueba final-. De esta manera, examinamos los cambios introducidos y si estos tendrían o no hacia una mejora en tanto que cumplían los criterios de la rúbrica. Este análisis se realizó a dos niveles, un primer nivel, en el que se examinaban las dimensiones de forma independiente y, un segundo nivel, en el que se relacionaban las dimensiones entre sí para considerar la prueba en su totalidad. La propia rúbrica nos permitió cuantificar los cambios.

En la Tabla 22 mostramos una relación de las variables analizadas y los indicadores utilizados.

Tabla 22

Relación entre las dimensiones, las categorías y los indicadores de análisis del estudio

Dimensión	Variable	Indicadores
Categoría GAPPISA	Competencialidad	Puntuación obtenida en (0-1 punto por apartado): 1. Característica situación inicial auténtica 2. Característica actividades auténticas 3. Característica complejidad cognitiva y emocional 4. Característica condiciones de resolución de la prueba
	Inclusividad	Puntuación obtenida en (0-1 punto por apartado): 5. Característica formato heterogéneo de las preguntas. 6. Característica formato heterogéneo de las respuestas.

		7. Característica atención a la diversidad.
	Potencial formativo	Puntuación obtenida en (0-1 punto por apartado): 8. Característica evaluación del proceso. 9. Característica criterios de corrección compartidos. 10. Característica comunicación formativa de los resultados.
Variables docentes	Curso	Todos los cursos que se han visto representados agrupados en: Educación Infantil, Educación Primaria y Educación Secundaria Obligatoria.
	Asignatura	Matemáticas, Lengua, Lengua extranjera, Ciencias Naturales, Ciencias Sociales.
	Años de experiencia docente	Agrupado en dos rangos: rango de menos de 15 años de experiencia docente y rango de más de 15 años de experiencia docente.
Variables formativas	Retroalimentaciones	Agrupado en número de retroalimentaciones recibidas por los formadores (de 1 a 3)

Obtuvimos puntuaciones en cada una de las versiones de las pruebas y con respecto a las tres variables de la Guía GAPPISA valoradas: nivel de competencialidad, nivel de potencial formativo y nivel de inclusividad. Los datos fueron tratados con el programa SPSS, versión 19, aplicando concretamente las pruebas de Correlación de Pearson, Anova y prueba de muestras independientes - Prueba de Leven y Prueba T-. De esta forma, comparamos las puntuaciones interpruebas y en relación con las variables docentes para establecer la significatividad estadística de los cambios.

2.3. Resultados

En este apartado vamos a exponer los resultados para poder dar respuesta a los objetivos planteados. Por un lado, vamos a describir los resultados cuantitativos, analizados estadísticamente, para explicar en qué nivel y de qué manera han mejorado las pruebas. Por otro lado, vamos a exponer los tipos de cambios que se han aplicado en la autenticación de las pruebas y a describir algunos ejemplos de diversas materias y niveles para cada una de las categorías de la Guía GAPPISA.

2.3.1. Análisis de las mejoras en las pruebas de evaluación.

A nivel de resultados cuantitativos, los datos que hemos analizado han sido: primero, la puntuación final obtenida por los docentes en cada una de las pruebas que han realizado contando las tres versiones (prueba original, prueba autenticada y prueba final); segundo, la puntuación por cada una de las características que conforman la Guía GAPPISA en cada una de sus versiones; y, tercero, la puntuación total en cada una de las variables de competencialidad, potencial formativo e inclusividad -obtenida a través de sumar el total de las características que se corresponden con la variable-. Seguidamente se han analizado los cambios en las pruebas según las diferentes variables recogidas que podrían influir en el resultado, estas son, respecto a los docentes: materia, curso y años de experiencia docente, y respecto a la formación, el número de retroalimentaciones recibidas por parte del equipo de formadores.

Antes de describir las pruebas y las correlaciones establecidas, es preciso señalar cual ha sido el nivel de participación en las actividades y el grado de implicación. Como vemos en la Tabla 23, de 206 inscripciones, 196 siguieron con el proyecto hasta el final. Además, 152 realizaron la segunda entrega lo que nos demuestra que los docentes estaban motivados para seguir aprendiendo y valoraban positivamente la retroalimentación aportada por los asesores.

Tabla 23

Número de entregas de los docentes según la versión de la prueba

	Prueba Original	Prueba Autenticada	Prueba Final
Número de entregas	174	196	152

Veremos a continuación los resultados agrupados en estos dos grandes bloques: cambios en las pruebas y relación con otras variables.

2.3.1.1. Cambios en las pruebas.

En primer lugar, en relación con las diferentes versiones de las pruebas -Versión Original, Versión Autenticada y Versión Final-, como se puede observar en la

Tabla 24, la puntuación media nos muestra una clara mejoría en las mismas, con más de tres puntos de diferencia entre la Versión Original y la Versión Final.

Tabla 24

Datos estadísticos de la puntuación final de todas las pruebas en cada una de sus versiones

	Muestra	Mínimo	Máximo	Media	Desviación típica	Varianza
Puntuación de la prueba en su Versión Original	175	0	6	1,09	0,803	0,645
Puntuación de la prueba en su Versión Autenticada	196	0	6	2,81	1,115	1,242
Puntuación de la prueba en su Versión Final	153	0	7	4,52	1,603	2,569

Con el análisis estadístico entre las puntuaciones de las variables de la Guía GAPPISA para cada una de las versiones comprobamos también que cada vez se diseñan pruebas de evaluación más auténticas en tanto que se van incluyendo y/o mejorando las diferentes características que la conforman y que dicha mejora es progresiva. Hemos obtenido una correlación positiva entre la prueba original y la autenticada, es decir, a mayor puntuación en la prueba original, mayor puntuación en la prueba auténtica, $r=0.153$, $p=0,028$. Existe también una correlación positiva entre la prueba auténtica y la final, es decir, aquellos que tienen mayor puntuación en la prueba autenticada, tendrán mayor puntuación en la final, $r= 0.355$, $p<0.0001$.

En segundo lugar, pasamos a observar, de entre las características que componen la Guía GAPPISA, cuáles han cambiado más, calculando la diferencia entre la puntuación de la versión original y nivel de la prueba final. Constatamos que los cambios que han tenido más éxito se refieren a: la situación inicial auténtica, las actividades realistas y la complejidad cognitiva. Por el contrario, la evaluación del proceso y la inclusión de condiciones de resolución realistas, aspectos también esenciales en una evaluación auténtica, no han obtenido una mejora tan elevada. Además, la relación entre todas las características en la versión final de la prueba es significativa, dándose una asociación lineal positiva. En el momento en que se

cambia una característica aumenta la puntuación positiva del resto, es decir, que realizar un cambio no va en detrimento del resto de características, sino que por el contrario todas ellas son interdependientes.

Tercero, el análisis de la puntuación obtenida en cada una de las variables de competencialidad, poder formativo e inclusividad, nos muestra que hay un aumento significativo de cada una de ellas (ver Figura 3). La variable de competencialidad supera con creces al resto, todo y que partía de un nivel más bajo de puntuación en la prueba original. Como hemos descrito anteriormente, las características que más cambiaron se correspondían con esta variable.

En la variable de potencial formativo, comprobamos que ha habido una mejoría pero que los niveles alcanzados, son moderados. Esta característica de las pruebas de evaluación auténticas se relaciona con cambios asociados a posibilitar la autoevaluación y la coevaluación, a compartir los criterios de calidad, explicitar el proceso de resolución seguido y/o enfrentar situaciones imprevistas.

En cambio, vemos que, respecto a la variable de inclusividad, es decir, el hecho de realizar pruebas ajustada a los diferentes perfiles de aprendizaje no ha habido tanta mejoría. Falta, por lo tanto, un trabajo relacionado con variar las formas de presentación de la información y de comunicación de los resultados, así como valorar las ayudas relacionadas. Además, en esta variable ya se observan niveles iniciales, levemente superiores al resto, en el tipo de preguntas y respuestas, relacionadas con la atención a la diversidad, aunque dicha variable se refiere a tareas tradicionales como incluir una imagen o pedir la realización de un esquema.

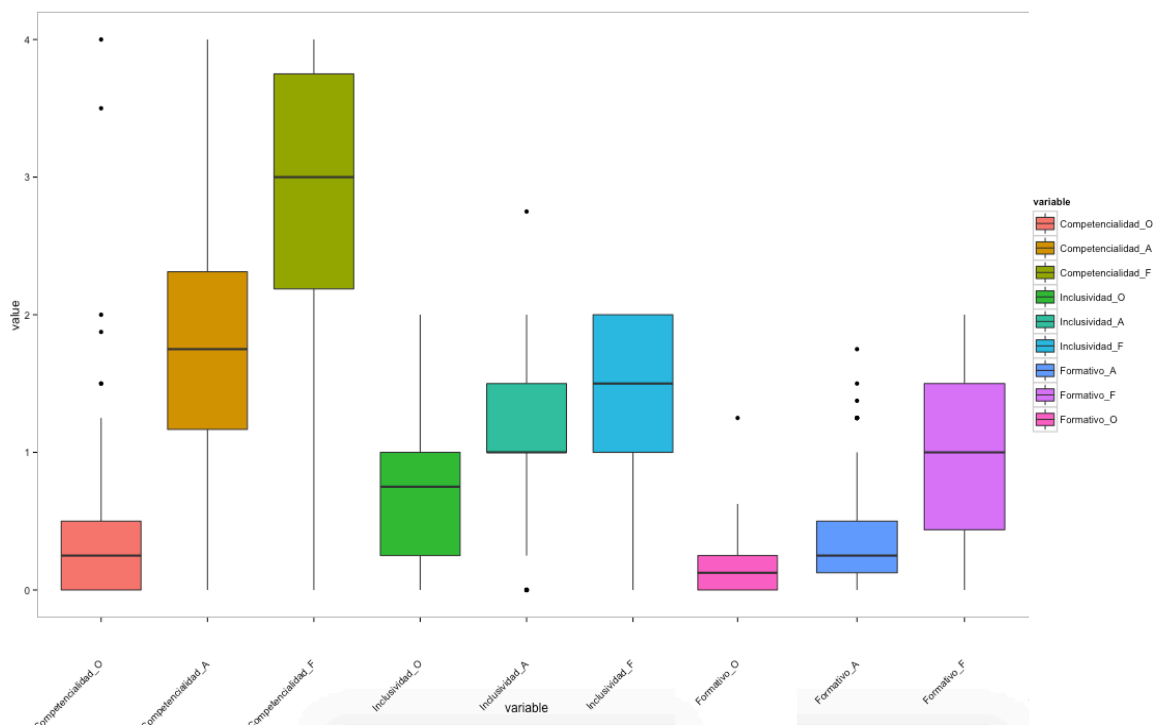


Figura 3. Nivel alcanzado en cada una de las pruebas según las categorías de competencialidad, poder formativo e inclusividad.

En resumen, los resultados estadísticos son significativos en las relaciones entre las variables referidas a las pruebas de evaluación -versiones, características y categorías-, es decir, los docentes han elaborado pruebas y/o actividades de evaluación más competenciales, formativas e inclusivas. A continuación, analizamos si existieron factores externos que interfirieron en esta mejora experimentada.

2.3.1.2. Relación con otras variables.

Primeramente, pasamos a analizar si las puntuaciones en cada una de las características de la Guía GAPPISA y de las tres variables, correlacionan y de qué forma con las variables referidas a los docentes como son: materia, curso y años de experiencia docente.

Respecto a la puntuación total en cada una de las características, no encontramos diferencias significativas con respecto a la materia y el curso. Para la variable años de experiencia docente, no se observa asociación lineal entre la puntuación total y los años de experiencia en dos rangos de edad -experiencia menor a 15,

experiencia mayor a 15-. Sí que existen diferencias significativas en las medias de la puntuación final en la característica de actividades auténticas: aquellos que tienen una experiencia menor de 15 años tienen puntuaciones superiores en actividades auténticas, $F=8.762$, $p=0.002$.

Considerando ahora la puntuación total de cada variable no encontramos correlación positiva con el curso y los años de experiencia docente. Sin embargo, sí hay diferencias significativas respecto a la asignatura en relación con las materias de lengua castellana y lengua extranjera y con la variable de inclusividad. Las pruebas de lengua castellana son menos inclusivas a diferencia de las pruebas de lengua extranjera que lo son más, seguidas de las de ciencias naturales. Aunque es destacable este dato no nos parece un resultado que influya negativamente en la interpretación general, puesto que únicamente sucede para una de las variables.

Seguidamente, relacionamos la puntuación final con el número de retroalimentaciones recibidas por parte del equipo de asesores, y vemos que es positiva: a mayor puntuación final mayor retroalimentación -a un nivel de significancia $p<0,001$ -. Es decir, aquellos docentes que obtienen mayor puntuación en actividades auténticas son los mismos que tuvieron un mayor número de retroalimentaciones, $r=0,782$; $p<0,001$. Las características más afectadas positivamente fueron: actividades auténticas y complejidad cognitiva. La evidencia de que las dos características que más han mejorado, en relación con la retroalimentación recibida, se correspondan con la variable de competencialidad, corrobora el hecho de que dicha variable haya sido la que más ha mejorado de forma global.

2.3.2. Descripción de las mejoras en las pruebas de evaluación.

En este apartado vamos a describir los cambios realizados en las pruebas de evaluación elaboradas por los docentes y a mostrar algunos ejemplos. Organizaremos los resultados agrupándolos en cada una de las categorías de la Guía GAPPISA, para describir las mejoras introducidas gracias a la formación en cada una de ellas. Y, por último, hablaremos de aquellos cambios que los docentes

realizaron bajo una incorrecta interpretación de los criterios de calidad y que, por lo tanto, no llegaron a ser mejoras; los denominamos las *pseudomejoras* de evaluación auténtica.

Hemos escogido ejemplos de diferentes ediciones y que son representativos de cómo han mejorado las pruebas. Mostrando solo las características que conforman una de las categorías, de forma independiente. Sin embargo, como hemos comentado anteriormente, es preciso recordar que existe una interrelación entre las características más allá de cada categoría y que, además, el análisis siempre se realizó en comparación con la prueba original para poder apreciar los cambios que se han llevado a cabo y situarlos en contexto.

2.3.2.1. Las mejoras a nivel de competencialidad

Dentro de la variable de competencialidad, la característica que ha sido más integrada por los docentes en sus pruebas es la de situación inicial contextualizada. Todas incluían un primer apartado donde se narraba y planteaba un relato similar a algún problema realista e incluían datos y materiales de consulta reales -catálogo de precios, programa de horarios, artículo en su formato original, medición *in situ* de estructuras, etc.-.

De manera generalizada, tras ese relato, se presentaban una serie de tareas o ejercicios también realistas. Para conseguir que lo fueran, los docentes hacían dos tipos de cambios mayoritariamente: uno, enlazar los ejercicios siguientes con la situación planteada inicialmente de forma que condujeran a la solución -pautando los pasos a realizar, uno por tarea- y, dos, redactar el enunciado de los ejercicios de forma personalizada -utilizando el "Imagina que tienes que..." y añadiendo información rea-.

Por último, otro de los cambios, aunque menos significativos, consistía en que las tareas fueran progresivamente más complejas, siempre interdependientes entre sí, de forma que la última tarea era la más difícil e integraba el resto -normalmente consistía en la elaboración de un producto-.

Un ejemplo de es el siguiente: una prueba de 2º de Educación Secundaria del área de matemáticas, específicamente de la unidad de geometría, titulada "El Ayuntamiento nos necesita". La prueba original consistía en la elaboración de dos ejercicios de cálculo y dos problemas de geometría, descontextualizados, sin ninguna relación entre ellos y con datos poco realistas y relevantes para los estudiantes.

Después del uso de la Guía GAPPISA y las orientaciones recibidas durante la formación y la retroalimentación de los asesores, el docente elaboró la siguiente prueba (ver Figura 4). La prueba empieza por una situación inicial bajo la demanda de "el ayuntamiento te pide asesoramiento como arquitecto", en la que se describen todas las condiciones para resolver el encargo. A ésta le siguen actividades realistas y relevantes, como son: la resolución geométrica del problema considerando el objetivo, debatir entre diversas opiniones sobre cómo resolverlo, la elaboración de un posible recorrido, etc. Dichas tareas van creciendo en complejidad de forma que los estudiantes empiezan por analizar los datos, para pasar a interpretarlos y comunicarlos en un informe científico-matemático final. Además, una de las tareas se resuelve en grupo, tal y como se haría en un contexto profesional trabajando en equipo.

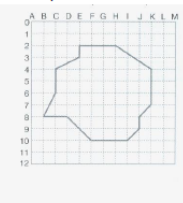
<p>"El ayuntamiento nos necesita"</p> <p>Imagina que recibes una llamada del Ayuntamiento de Tafalla porque quiere realizar dos nuevas calles. En la reunión con el concejal te explican lo siguiente:</p> <p>El siguiente fragmento está extraído del Acta de la Oficina Técnica del Ayuntamiento de la ciudad de Tafalla.</p> <p>Se decide que se realizarán dos nuevas calles en el barrio "Las Torres" (el área delimitada por el polígono en el mapa adjunto a escala 1:10.000) denominadas "travesía N-S" y "travesía E-O" respetando los siguientes vínculos:</p>  <p>Art. 1 : Las dos calles tendrán un recorrido rectilíneo.</p> <p>Art. 2: La travesía N-S deberá unir un punto del lado de vértice E2 y H2 con un punto del lado de los extremos F 10 e I 10 del polígono que delimita el barrio, la travesía E-O deberá unir un punto de vértice C4 y C6 con un punto del lado de extremos K4 y K7.</p> <p>Art.3: Las dos calles se intersectan en ángulo recto.</p> <p>Art.4: Las cuatro zonas en las que se dividirá el barrio por las dos calles deberán ser entre ellas equivalentes, es decir, tendrán la misma área.</p> <p>1) Asumiendo como unidad de medida el cuadrado de la retícula, ¿cuál es el valor del área S que ocupa el barrio? S = Cuadrados Luis opina que como se pierden zonas debido a las esquinas, no podemos considerar que dos medios cuadrados equivalen a uno entero ¿estás de acuerdo?</p>	<p>María cree que si utilizamos determinadas zonas como parques o zonas verdes mejoraremos mucho la habitabilidad de la calle. ¿Estás de acuerdo? ¿Cuáles utilizarías?</p> <p>2) TEST: ¿Cuánto mide en metros cuadrados la superficie real de cada una de las cuatro zonas en las que resultará dividido el barrio, asumiendo que los cuadrados de la retícula tienen 1 cm de lado?</p> <p>A) 1.300.000 m² B) 52.10³ m² C) 13.10³ m² D) 13.10⁴ m²</p> <p>3) Propón un recorrido para las dos calles que respete los vínculos impuestos en la deliberación municipal, asignando con precisión las coordenadas (letra, número) de los puntos por los que los recorridos atraviesan los confines del barrio.</p> <p>4) ¿Crees que la respuesta dada a la pregunta precedente es la única posible? Suponiendo que se puedan encontrar alternativas, ¿el largo de los recorridos resulta el mismo?. Examina brevemente la cuestión explicando tu razonamiento. En la siguiente sesión expondrás en clase tu punto de vista y lo defenderás frente a tus compañeros.</p> <p>5) Elabora un informe científico – matemático con las conclusiones principales.</p> <p>6) La siguiente pregunta se resolverá en clase en los grupos asignados y tendrá una puntuación del 20% de la nota: Si el polígono que delimita el barrio fuera un cuadrado, ¿cómo responderías a las preguntas anteriores? Propón la mejor estructura que debería tener el barrio para aprovechar lo mejor posible el espacio.</p>
---	--

Figura 4. Prueba competencial de matemáticas, de 2º ESO, después de la formación.

Así pues, vemos que es una prueba que destaca principalmente por su carácter competencial dado que permite a los estudiantes aplicar los conocimientos de geometría y las estrategias adquiridas en una situación con la que se pueden encontrar los estudiantes en el ámbito laboral, un problema que está descrito al inicio de la prueba. A continuación, se deben realizar ejercicios, que tienen sentido dentro de la propuesta inicial contextualizada puesto que se relacionan con ella de forma explícita. Por último, las últimas tareas son más complejas -elaboración de un informe y realización de otro caso similar en grupo- y deben dar cuenta de la competencia del estudiante

2.3.2.2. Las mejoras a nivel de poder formativo.

Para la variable de potencial formativo, los cambios más significativos estaban asociados a la característica de evaluación del proceso. La mayoría de las pruebas incluían alguna pregunta de reflexión sobre cómo se realizó alguna de las tareas o la prueba en su totalidad. En este ejercicio se preguntaba por el antes, el durante y el después de la prueba, de manera que los alumnos iniciaran un proceso de metacognición que los llevase a ser consciente del proceso y mejorar sus respuestas.



Otro de los cambios, consecuente con este proceso de reflexión, consiste en compartir los criterios de calidad -puntuaciones e indicadores de evaluación- que llevan a una resolución competente. Normalmente lo hacían para cada pregunta en el caso de la puntuación y mediante un listado de indicadores final o una rúbrica de evaluación. También en proponer alguna tarea de autoevaluación y coevaluación para mejorar la respuesta y aprender mientras se evalúa. En su conjunto, fueron en menor proporción los que optaron por estos tipos de mejoras.

Un ejemplo de prueba de calidad en esta categoría es el siguiente: una prueba de 3º de Educación Secundaria, del área de Lengua y Literatura, de la unidad didáctica que versa sobre la obra literaria de Miguel de Cervantes, *El ingenioso hidalgo don Quijote de la Mancha*. La prueba original que utilizaba el docente era muy poco competencial, solo evaluaba contenidos y no ofrecía ningún recurso ni indicadores que pudieran facilitar una reflexión y aprendizaje sobre el proceso. Consistía en

preguntas donde se pedía la definición de conceptos -qué son los libros de caballerías, qué son las novelas ejemplares, etc.- y una breve redacción sobre la biografía de Miguel de Cervantes.


Después de la formación recibida, el docente elaboró la siguiente prueba (ver Figura 5) en la que destaca su carácter formativo. Esta prueba mejorada propone tareas más competenciales como son investigar y redactar una noticia, y más auténticas, como es el hecho de que se deban realizar en equipo -tal y como podría ocurrir en la redacción de un diario-. Pero esta prueba destaca, especialmente, por las ayudas que ofrece para realizar estas tareas con éxito, como son: ofrecer recursos para conseguir información -direcciones web, palabras claves, etc.-, aportar técnicas para recoger y analizar dicha información -tabla para recoger datos y pauta de contenido- e indicadores de calidad para que se autoevalúen y puedan mejorar su producto final -rúbrica de proceso de cada una de las actividades-.

Miguel de Cervantes y *El ingenioso hidalgo don Quijote de la Mancha*

- Grupos de 4.** Recoged los datos biográficos más importantes de la vida de Miguel de Cervantes. Consultad al menos dos de las siguientes páginas. Redactad la respuesta en el anexo I. (10 puntos).
<http://www.quixote.tv/cerv.htm>
<http://www.biografiasyvidas.com/monografia/cervantes/>
<http://cervantes.uah.es/biografia/cronologia.htm#1558>
<http://cervantes.uah.es/biografia/cronologia.htm#1558>
<http://miguelde.cervantes.com/>
<http://www.youtube.com/watch?v=qtLaR2ftfwk>

- Grupos de 4.** Recoged información sobre la obra *El ingenioso hidalgo don Quijote de la Mancha*. Consultad al menos dos de las siguientes páginas. Redactad la respuesta en el anexo II. (10 puntos).
<http://www.auladeletras.net/material/cervan.pdf>
<http://www.rinconcastellano.com/renacimiento/quijote.html>
http://iris.cnice.mecd.es/lengua/profesores/eso3/t2/teoria_5.htm#IV
<http://www.youtube.com/watch?v=JZVZX8InGI0>

- Individual.** Redactar una noticia en la que se reproduzca uno de los siguientes capítulos del Quijote:
 1ª parte: III, VIII, IX, XVIII, XXII, XXXV
 2ª parte: X, XII

Lee la siguiente página e infórmate de cómo se redacta una noticia. Luego, completa el anexo III con la información que hayas recogido. (5 puntos) Lectura individual.
<http://www.icanito.cl/enciclopedia/articulo/segundo-ciclo-basico/lenguaje-y-comunicacion/lectura/2009/12/98-8783-9-texto-informativo-la-noticia.shtml>

Recoge información sobre el episodio que has leído. Anexo IV. Finalmente, redacta la noticia. Ten en cuenta la información del anexo III. (25 puntos).



	EXCELENTE	BIEN	MEJORABLE	DEFICIENTE	PUNTAJACIÓN
ACERCAMIENTO A LA VIDA DE CERVANTES (ANEXO I)	Todos los ítems están resueltos de forma correcta.	Resuelve de forma correcta 6 de los ítems. Las demás respuestas presentan algunos fallos.	Resuelve de forma correcta un 4 de los ítems. Las demás respuestas son incorrectas.	Se resuelven 2 o menos de los ítems de forma correcta. Las demás respuestas son incorrectas o no hay respuestas.	
SOBRE 10 PUNTOS	10	7	4	1	
EL LIBRO (ANEXO II)	Todos los ítems están resueltos de forma correcta.	Resuelve de forma correcta 4 de los ítems. Las demás respuestas	Resuelve de forma correcta 2 de los ítems. Las demás respuestas	Se resuelven menos de 2 ítems de forma correcta. Las demás	

PLANTILLA PARA EL CONTROL DE LA NOTICIA

En el titular

- ▣ Tienes que despertar el interés.
- ▣ Tienes que aportar información.
- ▣ Busca una frase con sentido completo.

Puede ser una oración afirmativa (verbo en presente) o frase

En la entrada

- ▣ Debes recoger la información básica de la noticia.
- ▣ Debes responder a qué, quién, cuándo, dónde, por qué y cómo.

En el cuerpo de la noticia

- ▣ Debes ordenar los datos según su importancia.

Figura 5. Fragmentos de la prueba competencial de lengua y literatura, 3º ESO, después de la formación.

Así pues, destaca su carácter formativo puesto que a través de las pautas e indicadores que se facilitan -y que en ningún caso son la solución al problema planteado-, el alumno puede resolver la prueba siendo consciente de su propio proceso y auto evaluándose para mejorar y, de esta manera, aprender.

Por último, cabe destacar la conexión entre las actividades competenciales propuestas en esta prueba y los recursos para guiar el proceso de resolución de esta. Este ejemplo nos enseña que, cuanto más competencial sea la propuesta inicial, más estrategias requieren los estudiantes para resolverla y, por lo tanto, en más oportunidades de aprendizaje participa. Es importante entonces, que también se pregunte a los estudiantes sobre el uso que han hecho de los instrumentos propuestos y evaluar si han sido o no estratégicos en su uso.

2.3.2.3. Las mejoras a nivel de inclusividad.

En la variable de inclusividad, la característica más integrada en las pruebas des de las pruebas iniciales como ya comentamos en los resultados previamente expuestos, fue la de diversidad en el formato de pregunta y respuesta. Esta característica aumentó también en cuanto los docentes incluyeron las primeras dimensiones relativas a la competencialidad. Y en este cambio, empezaron a emerger propuestas mucho más auténticas como presentar la información mediante un audio o vía online, y realizar demandas que incluían diferentes lenguajes.

También se encontraron algunas mejoras, aunque en una proporción muy baja, relacionadas con la característica de atención a la diversidad, como fueron aportar una hoja con ayudas que previamente se habían consensuado -fórmulas, esquemas, etc.- o la posibilidad de pedir ayuda a un compañero o al docente.

Un ejemplo de esta variable es la siguiente prueba de evaluación de 4º de Educación Secundaria, del área de Historia, de la unidad didáctica relativa a la II Guerra Mundial. La versión original que inicialmente utilizaba el docente era una prueba memorística donde se pedía la resolución de ejercicios como definir, decir datos y reflexionar sobre alguna cuestión; una prueba en definitiva que no permitía evaluar competencias y que con su estructura rígida y formato único -textual y de ejercicios cerrados- tampoco permitía atender a los diferentes perfiles de aprendizaje que haya en el aula.

Sin embargo, en la prueba final que realizó el docente (ver Figura 6), observamos que hay diversidad en el tipo de tareas como son elaborar un discurso, escribir un

guion, dramatizar una obra de teatro... Esta variedad en el tipo de tareas da la posibilidad a los estudiantes de expresar lo que saben de diferentes formas permitiendo que aquellos que tengan un lenguaje más desarrollado o ciertas habilidades que no son las tradicionalmente más valoradas, puedan comunicar sus aprendizajes. También, como las pruebas anteriores, contiene algunas actividades de carácter competencial como son una situación contextualizada -un historiador que trabaja en el museo- y tareas realistas - escribir una obra de teatro-, además de algunos criterios de evaluación que orientan la resolución de la prueba. Todo ello, en su conjunto, también permite una atención a la diversidad en tanto que aumenta la motivación y permite la autoevaluación.

<p style="text-align: center;">4º A ESO. HISTORIA. TEMA 9. LA II GUERRA MUNDIAL.</p> <p>NOMBRE _____ Nº _____</p> <p>FECHA _____</p> <ol style="list-style-type: none"> Durante el próximo verano eres contratado por un museo sobre la II Guerra Mundial. En él, eres el guía que debes explicar a los visitantes los diferentes acontecimientos del conflicto. El primer día de trabajo al llegar los primeros invitados debes comenzar con la visita. ¿Cómo comenzarías? <ol style="list-style-type: none"> Diciendo quién gana la guerra. Haciendo un número de circo. Explicando la batalla de Inglaterra Explicando cuales fueron las causas del conflicto. ¿Cómo lo harías? Escribe la información que les darías en esa presentación. Recuerda que durante este ejercicio se valorará lo siguiente: <ul style="list-style-type: none"> o El lenguaje utilizado (Siendo importante utilizar un lenguaje correcto, donde las palabras técnicas son importantes y las explicaciones deben estar debidamente unidas). o Una presentación adecuada (presentándose uno mismo y explicando cómo es el museo). o Realizar la localización histórica del conflicto y dando sus razones. o Un museo de historia, normalmente suele tener diferentes partes. La zona de documentos, la zona de material (restos de esa época de la historia), zona de audiovisuales (Mapas, videos explicativos, videos originales...) Lee atentamente el siguiente texto. <p style="font-size: small;"><i>"Esto es inhumano, ya no lo soporto más. El frío nos está matando: llevamos más de un mes sobre la nieve y el hielo, con temperaturas bajo cero, y la ropa que tenemos no es la adecuada.</i></p> 	<p style="text-align: center;">"TEATRO DE LA II GUERRA MUNDIAL"</p> <p>Durante este tema 9 hemos trabajado las causas, desarrollo, final y consecuencias de la II Guerra Mundial. Para comprender la evolución de la guerra divididos en grupos de 7 personas vais a realizar un teatro para representar el transcurso de la contienda.</p> <ul style="list-style-type: none"> • Este teatro tiene unas limitaciones, que son las siguientes: <ul style="list-style-type: none"> o Es un teatro mudo. Hay que escenificar los acontecimientos de la guerra usando únicamente la expresión corporal. NO SE PUEDE HABLAR. o Es posible que utilizando el papel se creen materiales (Barcos, aviones, bombas...) o Las onomatopeyas o reproducciones de sonidos también podrán utilizarse durante la representación. • Las pautas para realizar el teatro son las siguientes: <ul style="list-style-type: none"> o Cada alumno representará a un país. o Los países a representar son: Alemania, Italia, Japón, URSS, Estados Unidos, Francia, Gran Bretaña. o Durante la guerra hay multitud de acontecimientos, pero únicamente vamos a representar los siguientes: <ul style="list-style-type: none"> ▪ Creación de alianzas. ▪ Invasión aliada de Alemania. ▪ Rendición de Alemania. ▪ Bomba atómica. • Habrá que realizar para ello las siguientes actividades. <ul style="list-style-type: none"> o Creación de un guion <ul style="list-style-type: none"> ▪ El guion debe estar dividido en dos partes. En una se cita el hecho histórico y en el segundo se cita y explica vuestra representación. ▪ Ejemplo: Alemania declara la guerra a Francia. Hecho histórico. Alemania (Adrián) con el puño cerrado y el pulgar de la mano derecha extendido, se pasa dicha mano por el cuello mientras mira desafiante a Francia (Rubén). Representación o Representación del teatro.
--	---

Figura 6. Fragmentos de la prueba competencial de historia, 4º ESO, después de la formación.

Hay que destacar, por último, en relación con esta categoría, que las modificaciones para incluir aspectos que atiendan la diversidad debe ir precedida de un análisis de los perfiles de aprendizaje de los estudiantes, considerando características como las siguientes: concepciones que tienen sobre el aprendizaje y la evaluación, las estrategias que utilizan para aprender y de qué forma les gusta hacerlo y las emociones les genera y qué atribuyen sus éxitos y fracasos. Análisis que los profesores no hicieron.

2.3.3. Las *pseudomejoras* en las pruebas de evaluación.

Por último, del análisis de las pruebas de evaluación autenticadas surgió una nueva variable que denominamos *pseudomejoras*, refiriéndonos a aquellas mejoras introducidas por el docente que respondían a una interpretación errónea de las recomendaciones de cambio expuestas para cada característica de la Guía GAPPISA -los indicadores que conforman la categoría GAPPISA-. Y aunque no la contabilizamos como variable de análisis puesto que además ya quedaba reflejada en la puntuación no positiva, nos parece oportuno recogerla en este análisis cualitativo.

Las *pseudomejoras* que identificamos por cada una de las dimensiones son las siguientes:

1. Situación inicial contextualizada: proponer una situación realista que no aporta ningún contexto o que, realmente, no es una situación. Un ejemplo de ello es el siguiente enunciado inicial: "Imaginad que estáis en esta situación: el profesor os pide que realicéis el siguiente examen". Toda situación inicial debe estar bien contextualizada y, sobre todo, debe exponer los contenidos trabajados en la materia, si no, no tiene sentido y quedaría muy forzado o nada realista.
2. Actividades auténticas: llenar la prueba con fotografías de datos verídicos o apoyos visuales reales que no tienen nada que ver con la prueba. Si bien una recomendación es añadir información real para darle autenticidad a las tareas, también es cierto que estas deben aportar información que el alumnado deba utilizar. Si no es de esta manera, estamos añadiendo una cantidad de distractores innecesarios.
3. Complejidad cognitiva: bajar o subir excesivamente el nivel de la prueba. Un ejemplo de ello es pedir una actividad final que requiera el uso de estrategias que no han trabajado con anterioridad en ese curso, o dejar de incluir conocimientos esenciales de la disciplina. Por tanto, es importante que aunque se pidan tareas que son más complejas en cuanto elaboración y creación, estas no supongan una pérdida del trabajo de contenido estudiados en clase.

4. Formato heterogéneo de enunciados: modificarlos sin que esto suponga una atención a la diversidad. Un ejemplo de ello es cambiar el tipo de letra del enunciado o incluir imágenes que, en lugar de ayudar, dificultan su lectura y, además, no son muy realistas.
5. Diversidad en las modalidades de respuesta: proponer ejercicios variados pero todos textuales y de bajo nivel cognitivo. Por ejemplo, muchos docentes transformaban sus preguntas sobre conceptos en preguntas tipo test, de elección múltiple, de rellenar una frase, de relacionar, etc. En todos los casos, la demanda es la misma, una respuesta escrita, cuando lo que se debería requerir para atender a la diversidad es una variedad de respuestas en diferentes lenguajes (visual, corporal, etc.).
6. Condiciones de resolución de la prueba realistas: introducir incidentes críticos que nada tienen que ver con la contextualización de la prueba. Por ejemplo, en medio de una prueba de matemáticas decir "¡Alarma! simulacro de incendio"; o hacer ver que te llaman por teléfono y salir de la clase en mitad de la prueba y, a escondidas, vigilar si los alumnos copian. El objetivo de introducir un incidente crítico tiene que ver con recrear una posible eventualidad tal y como pasaría en la vida real cuando se estuviera trabajando en el contexto planteado (por ejemplo, estás redactando el artículo final para publicar en el blog y te llegan datos nuevos del estudio, igualmente debes entregarlo al día siguiente, ¿qué haces?)
7. Evaluación tanto del proceso como del resultado: preguntar por el proceso de resolución en una pregunta de bajo nivel cognitivo que únicamente pide reproducir conocimiento. No tiene sentido preguntar que has pensado antes, durante y después cuando únicamente se te pide que digas la fecha de nacimiento de Federico García Lorca.
8. Puntuaciones y criterios compartidos: explicitar únicamente las puntuaciones para cada pregunta o realizar rúbricas que no aportan indicadores. El primer caso, es muy común en las versiones originales, y es positivo que los alumnos sepan cuánta puntuación tiene cada tarea, para poder ser estratégicos resolviéndola, pero sin unos criterios cualitativos de corrección, esto no resulta una ayuda. En el segundo caso, un ejemplo de rúbrica incorrecta sería el siguiente: para el criterio "elaborar un mapa de

conceptos", poner tres niveles, no lo hace o lo hace mal, lo hace regular y lo realiza correctamente. En ninguno de los niveles se comparten los criterios puesto ¿qué significa hacerlo mal? ¿qué diferencia hay entre hacerlo regular o bien? La rúbrica debe servir para compartir los pasos a realizar y los indicadores de calidad de manera que, al leerla, el alumnado pueda corregirse.

9. Comunicación de los resultados: no hacer que la devolución de la corrección por parte del docente sea una oportunidad de aprendizaje. En este aspecto encontramos una lista variada de *pseudomejoras* como son: - preguntar por el proceso, pero no realizar una corrección cualitativa - aportar una rúbrica de evaluación, pero luego no permitir una sesión de reflexión para resolver dudas - realizar una devolución cualitativa pero no permitir que puedan llevarse la prueba (o una copia de esta) para repasarla, etc. Es importante entender que el proceso de aprendizaje (y también el de evaluación), no acaba cuando finaliza la prueba, sino cuando se efectúa la devolución y se reflexiona en base a ella.
10. Atención a la diversidad: asociar ayudar a los alumnos con necesidades con el hecho de rebajarles la dificultad de la prueba. Un ejemplo es repartir la prueba y, a unos alumnos en concretos, decirles "no hace falta que realicéis estas preguntas" y tachárselas. El objetivo de atender la diversidad no se puede traducir en no evaluar su actuación competencial, sino que es aportar las ayudas necesarias para que puedan conseguirlo (adaptando lenguaje, guiando la resolución, etc.). Además, debemos prestar especial interés a cómo se sienten los niños si se les trata diferente y se les deja en evidencia.

En definitiva, los resultados relativos a este punto suponen un buen diagnóstico para identificar aquellos aspectos que deberían clarificarse mejor antes de utilizar la guía. en este aspecto nos muestran qué aspectos deben aclararse mejor en el uso de la guía.

2.4. Discusión de los resultados

A continuación, discutiremos cómo los resultados dan respuesta al propósito del estudio: analizar los cambios en las pruebas de evaluación para determinar si ha habido una mejora significativa en el grado de competencialidad, carácter formativo e inclusividad de las pruebas elaboradas por los docentes tras la formación y valorar las fortalezas y debilidades de la formación GAPPISA como instrumento para promover los cambios señalados.

Después de analizar las pruebas de evaluación diseñadas por los participantes, podemos afirmar que los resultados de la formación con el uso de la guía GAPPISA son positivos dado que dichas pruebas son ahora más auténticas y ayudaron a los aprendices a desarrollar sus competencias, como ya apuntaron parcialmente estudios anteriores (Álvarez-Bernardez y Monereo, 2016; Monereo, Castelló et al., 2009; Monereo et al., 2012; Monereo y Morcillo, 2011). Estos cambios se reflejan tanto en la progresión positiva e interrelacionada durante la mejora de las pruebas, como en un aumento de su competencialidad, carácter formativo e inclusividad. Los cambios, que pasamos a discutir a continuación, constituyen un modelo de evaluación a seguir para la implementación de una evaluación auténtica en el aula, instrumentalizada por productos elaborados por los propios docentes.

Empezando por la variable de competencialidad, observamos que se han modificado todas las características que la conforman, especialmente aquellas que definen que la prueba sea más realista y relevante para los aprendices. En efecto, el primer aspecto que es preciso considerar para evaluar competencias es la presentación de un reto, de una situación problemática a resolver, situada necesariamente en un contexto auténtico (Ashford-Rowe, Herrington y Brown, 2014; Monereo, 2009). El uso de casos de enseñanza, exposiciones y problemas a resolver son métodos de evaluación auténtica que precisan todos de una contextualización (Darling-Hammond y Snyder, 2000). La autenticidad en dichas situaciones es la clave del desarrollo de las competencias, entendidas como el uso de los conocimientos y habilidades para resolver problemas complejos que un ciudadano puede encontrarse en diversos ámbitos.

Por otro lado, la introducción de una situación inicial contextualizada tiene un efecto "bola de nieve" con el resto de las dimensiones. Así pues, no basta con describir un problema, sino que de este emergen, de forma consecencial, mejoras en la tipología de tareas, las cuales pasan a ser más próximas a las que se darían en la vida real -alejadas de ejercicios como definir o responder preguntas sin un porqué- (Darling-Hammond y Snyder, 2000) y el hecho de resolver tareas intermedias sin perder de vista las metas -incremento de la complejidad- (Avery et al., 2012). Además, el problema y el realismo de las propuestas propician la participación en contextos grupales -tareas en trabajo cooperativo- y el uso de información real y de estrategias para controlar el proceso -guías, tablas con indicadores, bases de orientación, etc.- (A. Herrington y J. Herrington, 2006). Todos estos elementos se materializaban en la construcción de la prueba casi de forma automática cuando el docente comprendía que era necesario aplicar los conocimientos en una situación realista con un contexto y un proceso de resolución complejo.

Respecto a la variable de inclusividad, de forma mayoritaria comprobamos que los docentes realizaron adaptaciones de la prueba según las preferencias de aprendizaje de su grupo de aprendices que eran conocidas por ellos, para asegurar la participación de todos los alumnos, especialmente en sus formas de acceder y comunicar sus aprendizajes (Sala et al., 2013). Estas adaptaciones no consistieron en elaborar pruebas especiales para aquellos alumnos que presentaban mayores dificultades, sino en reconocer las ayudas -recursos, indicadores, tareas de progreso- que cada uno de los aprendices necesitaba para resolver la prueba. También, se expusieron las demandas en formato audio o vídeo para permitir que los estudiantes con una preferencia de input más visual conectaran más con la tarea y tuvieran una comprensión mayor de la misma. También se diseñaron algunas tareas cooperativas para aquellos estudiantes que prefieren trabajar en equipo o que requieren de un apoyo más específico debido a alguna necesidad educativa, respondiendo a la vez a unas condiciones de resolución realistas (Gulikers et al., 2004).

Las dificultades que surgieron al respecto y que no permitieron alcanzar niveles de mejora tan elevados como en la variable de competencialidad, se relacionan con el desconocimiento que tenían los docentes sobre la manera de aprender de sus

alumnos. Poco se sabe sobre la identidad del aprendiz, sobre sus concepciones, estrategias y sentimientos en torno al aprendizaje y conviene utilizar instrumentos para evaluarla y, con ello, ajustar los procesos de enseñanza (Morodo, 2017). Esta es una carencia bastante habitual, especialmente en enseñanza secundaria, que requiere un trabajo previo de observación y seguimiento sistemático de los alumnos desde su entrada al centro educativo.

Además, cabe destacar que estas medidas de atención a la diversidad están en línea con la legislación educativa al respecto, la Orden EDU/849/2010 del Ministerio de Educación de España, por la que se regula la ordenación de la educación del alumnado con necesidad de apoyo educativo. En ella se expone que el alumnado con necesidades de apoyo educativo participará en las actividades de su grupo de referencia y que se pondrán todos los medios necesarios para que así sea. En Catalunya, específicamente, a estas medidas se les denomina medidas de soporte universal (Generalitat de Catalunya, 2015) y son aquellas actuaciones que dan respuesta a la diversidad que se centran en cambiar las metodologías que afectan a todo el grupo clase -por ejemplo, aprendizaje cooperativo, evaluación formativa, etc.-, aspectos que también integramos en nuestra propuesta y que ha sido utilizados por los docentes.

Por último, en la variable relativa al carácter formativo de la evaluación, encontramos propuestas mayoritariamente enfocadas a que los alumnos tomaran consciencia sobre cómo resuelven los problemas, analizando los pasos seguidos y las evidencias para así poder actuar estratégicamente y aprender de sus errores (Mc.Laren, 2012). También, se implementaron algunas propuestas de autoevaluación y coevaluación, actividades que la literatura ha señalado como procesos reflexivos que permiten la autorregulación mediante la valoración de la consecución de unos criterios conocidos previamente (Panadero y Alonso-Tapia, 2013; Sanmartí, 2010b). Y, en menor medida, se compartieron criterios de evaluación con los alumnos, aspecto clave para poder aprender durante la evaluación (Dawson, 2015).

Sin embargo, esta variable no ha alcanzado niveles muy elevados. Las razones que lo explican son de diversa índole: desde la falta de habilidades metacognitivas

del alumnado, fruto de la enorme carencia de propuestas didácticas orientadas en este sentido durante toda la escolaridad, hasta la inexistencia de un trabajo docente dirigido a dialogar con los alumnos sobre la evaluación, explicitar los objetivos de aprendizaje, practicar formas de autoevaluación y coevaluación o consensuar criterios de evaluación (Romero-Martín, Castejón-Oliva, López-Pastor y Fraile-Aranda, 2017). De hecho, según Brown y Harris (2014) hay que integrar la autoevaluación como contenido de estudio, antes de como un método de evaluación, puesto que para que sea efectiva los alumnos primeramente deben desarrollar esta competencia.

Pasamos ahora a discutir las implicaciones que ha tenido el contexto para la mejora de las pruebas. En relación con las variables que podrían modificar la calidad de las pruebas, resolvemos que ninguna de las variables docentes, como el tipo de materia, el curso en el que se imparte docencia o los años de experiencia docente ha influido significativamente sobre los resultados. Consideramos que este dato es positivo en tanto que indica que la evaluación de una actuación competente, el ajuste de ayudas a los perfiles de aprendizaje o el hecho de considerar la evaluación como una oportunidad de aprendizaje son objetivos que pueden lograrse en la inmensa mayoría de contextos educativos.

Sí que ha habido una excepción en este aspecto, dado que las pruebas de lengua extranjera eran significativamente más inclusivas. El aspecto de inclusividad que más destacaba era el de formato heterogéneo de preguntas y respuesta. Seguramente este hecho es debido a que la estructura de enseñanza de la lengua extranjera ya se organiza a partir de diversidad de tareas agrupadas según las dimensiones de la competencia en comunicación lingüística, como son la comprensión lectora, la expresión escrita y la comunicación oral. Esta división responde a que esta área se centra claramente en promover competencias en los alumnos, en que puedan usar la lengua en diferentes contextos, huyendo así del objetivo de ser expertos en conocimientos de la materia, como sucede en otras áreas.

Esto nos indica que el objetivo de enseñanza último, el para qué y por qué enseñamos, qué propósito hay detrás, condiciona enormemente toda la didáctica

y, en especial, la evaluación. Y lo que determina una evaluación auténtica de los aprendizajes no es el área de estudio y sus contenidos, la edad y capacidades de los alumnos o la experiencia del docente titular, sino la relevancia y socialización de la práctica evaluativa y su poder para fomentar el aprendizaje de forma ajustada a los aprendices. Una muestra de ello es que cuando hablamos de evaluación de los aprendizajes encontramos experiencias exitosas en todo tipo de niveles educativos y disciplinas (López-Pastor y Pérez, 2017) y lo mismo cuando hablamos de evaluación auténtica (Frey et al., 2012; Fuentes y López, 2017; Koh et al., 2012; Martínez y Cantero, 2014; Meyers y Nutly, 2009; Raymond et al., 2013).

Para acabar la valoración de las mejoras en las pruebas, es preciso destacar nuevamente, que el análisis de estas debe de hacerse de forma global puesto que todas las categorías están interrelacionadas. Esto nos conduce a afirmar que podemos validar la consistencia teórica del instrumento Guía GAPPISA en el momento en el que observamos que las dimensiones se complementan y hay una correlación positiva, es decir, que modificar una característica provoca que otra tienda a mejorar. De esta forma, vemos como el hecho de diseñar pruebas más competenciales se complementa necesariamente con el diseño de pruebas más inclusivas y formativas. Hallamos pues, una consistencia en el marco teórico sobre las categorías que definen una evaluación auténtica.

Este resultado también se verifica con el hecho de que algunos docentes transformaron sus pruebas en proyectos de aprendizaje basados en la resolución de situaciones realistas, en algunos casos reales -como pintar el patio del colegio, rediseñar la biblioteca, construir y gestionar el huerto escolar, etc.-, mediante los que el alumnado podía demostrar sus competencias y en acción para una valoración *in situ*. Estos proyectos rompían los esquemas tradicionales del concepto de prueba como instrumento acotado en el tiempo, espacio y recursos, pero, sin embargo, respondían a la necesidad de diversificar los instrumentos y escenarios de evaluación, posibilitando la experimentación, las resoluciones alternativas y las potencialidades de los alumnos, un aspecto clave de la competencia en evaluación del docente (Tonda y Medina, 2013).

Valorando ahora la efectividad del proyecto de formación GAPPISA, primero es preciso recapitular el éxito de sus resultados y, segundo, las características de este que lo han hecho posible. En primer lugar, la correlación positiva entre las pruebas originales y las finales nos confirman que los docentes van avanzando de forma progresiva. La correlación positiva entre las puntuaciones en la prueba auténtica y la prueba final muestra que los docentes que han realizado un trabajo mayor en la primera versión han interiorizado mejor los aprendizajes de forma que en la segunda entrega han obtenido puntuaciones finales más altas con mayor rapidez. La diferencia entre la prueba autenticada y la final es el *feedback* del asesor.

En segundo lugar, creemos que dos de las características han sido claves: partir de las pruebas de evaluación de los docentes y realizar retroalimentaciones personalizadas que acompañen el proceso de mejora. Así es, una de las fortalezas del proyecto es que los docentes trabajen desde sus propias pruebas. Podríamos haber propuesto el diseño de carpetas de aprendizaje o de rúbricas de evaluación, pero estas prácticas no estaban en su zona de desarrollo. Existen experiencias que también apuestan por la transformación de los exámenes escritos como medida de implementación de la evaluación auténtica (Villarroel, D. Bruna, Bustos, C. Bruna y Márquez, 2018).

Por otra parte, también el profesorado debe implicarse en su aprendizaje y no hay mejor manera que conseguirlo que trabajando en el diseño de un instrumento que le reportará una mejora en su actividad docente. A la vez, el *feedback* recibido por los asesores, se realizaba no solo en comparación con los criterios de calidad expuestos, sino también con el producto del que partía el docente, su prueba. De esta manera nos situamos en la zona de desarrollo próximo del docente y, además de sus resultados, valorábamos el proceso. Esta mediación por parte del equipo asesor es otra fortaleza de la formación, comprobada en tanto que los resultados nos indican que, a mayor retroalimentación recibida, mejores resultados obtenidos. Además, el uso de otros instrumentos como la Guía GAPPISA, los materiales online, la base de pruebas, etc., que también guiaban la adquisición de niveles progresivamente más elevados de conocimiento aplicado, fue decisiva.

Nuevamente, también el hecho de acompañar al docente durante la aplicación de la prueba fue una actividad determinante, especialmente para mantener la motivación e implicación de los docentes. Un cambio en las prácticas de evaluación pasa por dar la oportunidad a los profesores de experimentarlo mientras reciben apoyo constante para transformarlo en una experiencia de aprendizaje, como ha sido el caso del proyecto GAPPISA. Este aspecto está en consonancia la visión del formador como asesor que acompaña a transformar la práctica educativa trabajando con los docentes, yendo más allá de impartir una formación únicamente (Imbernon, 2007; Sondergeld et al., 2010) y actuando como generadores de contextos que consiguen hacer emerger las condiciones favorables para la innovación educativa (Badia y Monereo, 2004; William et al., 2004).

Hay experiencias previas como la de Álvarez-Bernárdez y Monereo (2016) que, de forma cercana, también apoyan el protagonismo de los docentes como medio potenciador del aprendizaje, defendiendo que el modelo de investigación colaborativa convierte a los docentes en agentes activos y, en consecuencia, aumenta la posibilidad de que los cambios que aplican se extiendan en el tiempo. En este sentido, como decíamos, ha sido una debilidad no poder ofrecer un acompañamiento más dialogado al no contar con recursos suficientes como para acompañar a los orientadores del centro o disponer de uno en cada escuela. Sin embargo, quisimos solventar esta distancia introduciendo la figura del coordinador, como agente activo del cambio, que a su vez hiciera de asesor.

Por ello, también valoramos positivamente que, con la estructura del Proyecto GAPPISA, se cumple la intención de promover la construcción de una comunidad de práctica dentro de cada centro, en la que se puedan inscribir nuevos docentes y participar de la transformación elaborando pruebas auténticas, con la mediación motivacional de sus compañeros y los resultados que estos han obtenido. Este nos parece un muy buen inicio para convertir los centros en "comunidades de aprendizaje profesionales", uno de los factores determinantes para conseguir la sostenibilidad de las innovaciones educativas (Liesa et al., 2018). El hecho de que las inscripciones se mantuvieran da cuenta de la valoración positiva que hicieron los docentes que supieron generar la necesidad de cambio en algunos de sus iguales de claustro. Sin embargo, es necesario seguir investigando otras medidas

para que la implicación hacia el cambio se extienda a todo el equipo docente, con estrategias que se han demostrado como efectivas como por ejemplo el “coaching” entre docentes (Johnson et al., 2017; Kraft et al, 2018), la co-docencia o docencia compartida (Oller et al., 2018) o la figura del coordinador como supervisor interno, experto en competencias en evaluación (Moreno, 2015).

Por último, nos parece importante hacer mención al carácter semipresencial de la formación y al uso de la plataforma virtual, como posible factor influyente en la calidad de la formación. Como se ha demostrado, especialmente en el ámbito universitario, los entornos virtuales de aprendizaje son una buena estrategia para, no solo facilitar el acceso a la información, sino también para favorecer la interacción entre los aprendices y la socialización de los aprendizajes (Blasco, Mengual-Andrés y Roig-Vila, 2011). En este sentido, es relevante considerar esta variable en futuras investigaciones y plantear una evaluación de la calidad del entorno virtual de aprendizaje, basándonos en experiencias de evaluación de cursos en línea (Mengual-Andrés, Lloret y Roig-Vila, 2015); y, también, la actualización de este entorno mediante estrategias como el relato digital, que permite presentar los contenidos de forma atrayente y motivadora para los aprendices (Roig-Vila y Rosales-Stalkus, 2016).

2.5. Conclusiones, Limitaciones y Prospectiva

Concluimos pues, que la formación fue significativa en tanto que más de doscientos docentes diseñaron pruebas de evaluación más auténticas. Este proceso de cambio basado en partir de las pruebas tradicionales de evaluación enfocadas al análisis de lo que se sabe, a nivel de conocimientos, para llegar a actividades de evaluación centradas en la valoración de las competencias alcanzadas, mediante las que se promueve el aprendizaje de los alumnos desde la personalización de dichas prácticas, fue exitoso. Lo que nos mueve a contemplar que las características de este enfoque de evaluación son las que se deben de extrapolar a las prácticas de enseñanza si lo que se pretende es que los aprendices sean autónomos, proactivos y competentes ante las situaciones reales en las que se encuentren en los futuros contextos educativos, personales y sociales en los que tengan que participar a lo largo de su vida.

A nivel de limitaciones del estudio, entendemos que nos hemos centrado en las características pedagógicas de la formación y los resultados de esta a nivel de cambios en las pruebas de evaluación. Pero, lo que no hemos considerado es la incidencia que la aplicación de dichas pruebas tuvo en el aprendizaje de los alumnos, es decir, el resultado de la aplicación de la innovación en evaluación. Tampoco hemos considerado la voz del alumnado en el diseño ni la aplicación de las pruebas de forma directa. Nos parece que en estudios posteriores sería importante investigar estas dos cuestiones para describir con mayor profundidad los efectos que han tenido las pruebas en los aprendizajes considerando la voz de los alumnos, como ya se ha hecho en estudios anteriores en los que los estudiantes también opinaban sobre la calidad y la cantidad de las tareas de evaluación auténticas (Gulikers et al., 2006; Opazo, Pérez-Cabaní, y Sepúlveda, 2016; Opazo, Sepúlveda y Pérez-Cabaní, 2015).

Estrechamente relacionado con esta limitación, encontramos que únicamente nos hemos centrado en un instrumento de evaluación, justificado en el hecho de que este era el más utilizado por los docentes. Sin embargo, las dificultades de los profesores van más allá e incluyen todo el proceso de evaluación, desde la programación de esta, hasta la evaluación formativa a lo largo de la unidad. En este sentido, sería interesante diseñar también una guía que permitiera a los docentes reflexionar y mejorar su proceso de evaluación formativo, siguiendo el modelo planteado por Pérez-Cabaní y Carretero (2009) en el instrumento IRAESO, *Instrument de Reflexió sobre l'Avaluació de l'Aprenentatge*; una aplicación formativa que permite analizar situaciones sobre el proceso de evaluación que se han identificado como problemáticas para los docentes y describir las estrategias que se podrían utilizar para mejorar.

Por último, haciendo un poco de prospectiva, consideramos que hay dos planteamientos iniciales sobre los que podríamos seguir trabajando: por un lado, el impacto de la evaluación en la enseñanza de los docentes y, por otro lado, la permanencia de los cambios en evaluación; dos aspectos interrelacionados entre sí. Respecto a la primera cuestión, al ser la evaluación un aspecto tan influyente en la práctica educativa, creemos que muchos docentes se replantearon la forma de enseñar, bajo la premisa que modificar la evaluación obligaría a transformar los

métodos de enseñanza. Existen indicios de que las nuevas prácticas de evaluación desafían las actividades diarias de enseñanza y aprendizaje (Xie, Kyu, Sheng-Lun y Luthy, 2017) y que, por lo tanto, pueden cuestionar las bases pedagógicas.

Respecto a la segunda cuestión, nos preguntamos si los cambios en la evaluación serán sostenidos por los docentes a lo largo del tiempo de forma que se lleguen a integrar en su quehacer como docente. Investigar sobre si los cambios son sostenibles es otra manera de evaluar la influencia de la formación. Por lo tanto, nos parece muy relevante investigar si el proyecto GAPPISA impacta sobre la identidad profesional y, en especial, sobre la posición evaluadora dentro de la construcción como docente. Para medir el impacto, sería adecuado estudiar cuáles son las concepciones, prácticas y sentimientos que los docentes tienen respecto a la evaluación después de haber realizado la formación y pasado unos años. Con este cuestionamiento iniciaremos el segundo estudio.

Finalmente, concluyendo este apartado, estamos en condiciones de afirmar que la formación y el uso sistematizado de la Guía GAPPISA de manera contextualizada en cada realidad docente y, a la vez, de forma masiva en una red de escuelas que quieren modificar sus prácticas de evaluación, promueve una mejora sustancial en sus pruebas y esperamos que en futuros contextos lo podamos replicar con éxito e introducir las mejoras.

CAPÍTULO 3. Estudio 2: Cambios en el Posicionamiento Docente sobre Evaluación a partir de una Formación sobre Evaluación Auténtica

3.1. Introducción y Objetivos

De entre las distintas competencias que debe poseer un docente, una de las que destaca por su relevancia es la competencia en evaluación (Sanahuja y Sánchez-Tarazaga, 2018; Tonda y Medina, 2013). Conocemos la influencia que tiene la evaluación sobre el modo en que aprenden los alumnos y la forma en que enseñan los docentes (Morales, 2009; Tierney, 2006; Wiliam, Lee, Harrison y Black, 2004). Los efectos retroactivos de la evaluación sobre ambos procesos la convierten en uno de los principales mecanismos para el cambio y la innovación educativas. Por eso es necesario, que los docentes sean competentes en evaluación con el fin de asegurar la calidad educativa (DeLuca, 2012; Mellati y Khademi, 2018).

Además, estas aseveraciones se constituyen y desarrollan en el nuevo paradigma sobre enseñanza competencial. Ante las exigencias respecto a la formación de ciudadanos y profesionales competentes, capaces de afrontar los problemas de una sociedad en continuo cambio (Comisión Europea, 2007), las prácticas evaluativas tradicionales, centradas en una evaluación final y sumativa de conocimientos básicamente conceptuales y declarativos, están completamente obsoletas. La denominada “cultura del test” orientada a acreditar el aprendizaje mediante pruebas escritas, de opción múltiple y respuesta prioritariamente memorística, difícilmente puede resultar válida para atestiguar el dominio y aplicación de los conocimientos, estrategias y competencias que debe poner en juego un ciudadano del siglo XXI (Álvarez, 2001; Gómez y Miralles, 2012).

Sin embargo, estos exámenes en los que prima el conocimiento conceptual y memorístico siguen siendo el principal instrumento de evaluación en la enseñanza obligatoria, tanto en educación primaria (Gómez y Miralles, 2013; Calatayud, 2000) como en educación secundaria (Gómez et al., 2018; Monteagudo y Vera, 2017) en muchos centros españoles. Estos exámenes tradicionalmente nacen de concepciones sobre la evaluación que sostienen que su función es meramente

acreditativa y enfocada a medir el nivel de los conocimientos que los alumnos han adquirido. Esta visión es realmente insuficiente y supone desaprovechar toda la versatilidad de la evaluación.

Así pues, dar cuenta de los niveles de logro alcanzados por los alumnos es necesario dentro del sistema educativo, pero no es ni de lejos lo más significativo para que alumnos y docentes tomen conciencia de sus prácticas de enseñanza y aprendizaje y sepan qué deben mejorar y cómo hacerlo (Sanmartí, 2007). Es decir, para que la evaluación sea realmente educativa. Anclarse en definirla únicamente como un conjunto de calificaciones supone claramente un obstáculo para la calidad de la enseñanza, y provoca que los estudiantes no se impliquen con motivación en su aprendizaje (Gómez-Devís et al., 2017).

Es necesario, pues, invertir en una evaluación auténtica que permita evaluar competencias a través de la resolución de problemas realistas que signifiquen un reto, en los que el alumno tenga un rol participativo y se consideren sus necesidades y motivaciones (Frey et al., 2012; Monereo, 2009; Villarroel et al., 2017). Una evaluación de este tipo ha de servir para detectar expectativas, conocimientos y estrategias previas de los alumnos y adecuar de ese modo los métodos y contenidos. También debe ser una herramienta que posibilite seguir los progresos y dificultades de los alumnos de manera continua e in situ. Asimismo, debe ofrecer interesantes indicadores que permitan a los aprendices identificar sus debilidades y autorregular su aprendizaje, o a los docentes autorregular su enseñanza y remediar ciertas carencias detectadas. Incluso ser un elemento de cohesión curricular entre las distintas áreas a través de pruebas y proyectos interdisciplinares que todo ciudadano o profesional competente deba ser capaz de resolver.

A esta necesidad de centrarse en lo que Coll y Remesal (2009) han denominado el polo pedagógico de la evaluación, se le suma, como avanzábamos anteriormente, la necesidad de establecer una coherencia entre la enseñanza y la evaluación como caras de una misma moneda. Sabemos que implementar una nueva propuesta metodológica sin modificar la evaluación, genera resistencias y desconcierto en la comunidad educativa (Moreno, 2017). Con vistas de suplir esta necesidad, en la

última década, alentada por currículums educativos basados en enfoques competencias y en una educación inclusiva, ha habido un incremento exponencial de las solicitudes de formación del profesorado sobre nuevos instrumentos de evaluación.

Pese a todo, no parece que esa exigencia esté dando sus frutos en el sentido de promover alternativas de formación inicial y permanente que resulten suficientemente eficaces. La explicación debe buscarse en, por lo menos, dos aspectos. En primer lugar, parece que las prácticas evaluativas en las que participan los futuros profesionales de la educación –maestros, pedagogos, psicólogos de la educación, etc.- no han cambiado sustancialmente y los profesores universitarios, hablando en términos generales, no predicán con el ejemplo (Clark y Rust, 2006; Ion y Cano, 2012; López-Pastor y Pérez-Pueyo, 2017). Es decir, no evalúan mediante situaciones realistas y relevantes que obliguen a los estudiantes a afrontar problemas que se encontrarán en sus futuros contextos de desarrollo, aunque sí existen algunas experiencias positivas al respecto (Mejías y Monereo, 2016; Cano et al., 2012). Por otro lado, curricularmente el aprendizaje de la competencia en evaluación de los futuros maestros no tiene la suficiente importancia que debería tener dentro de los planes de estudio (Martínez, 2013; Popham, 2009; Sanahuja y Sánchez-Tarazaga, 2018).

En segundo lugar, los actuales modelos de formación permanente del profesorado sobre evaluación siguen mayoritariamente anclados en ofrecer conocimientos y estrategias puntuales sobre evaluación, que, si bien pueden traducirse en cambios en algunas actividades y prácticas de aula, también puntuales, no suponen un cambio realmente transformador, profundo, un cambio integral y radical de la identidad del docente. La mayoría de los estudios, aunque concluyen con resultados positivos en lo que se refiere a la mejora de la competencia en evaluación de los docentes, también señalan que podrían haber modificado más notablemente sus creencias y que deben continuar su proceso de aprendizaje (Abdul et al., 2009; Badia y Chumpitaz-Campos, 2018; Monereo, Castelló et al., 2009; Rodríguez, 2010; Vergara, 2011).

Ante este panorama, es imperante seguir investigando sobre cómo mejorar la competencia en evaluación de los docentes. En primer lugar, defendemos que es necesario intervenir en las concepciones que los docentes tienen, lo que Beijaard, Verloop y Vermjunt (2000) llaman identidad docente, para generar nuevas posiciones respecto a la evaluación (Monereo y Pozo, 2011), puesto que si estos no modifican sus creencias y sentimientos es muy probable que las nuevas prácticas que implementan no sean sostenibles (Monereo, Badia et al., 2009).

El cambio al que nos referimos requiere un nuevo modo de posicionarse frente a la evaluación no solo de hacer evaluación, es decir, una transformación en la manera de concebir sus funciones y finalidades, de desarrollar nuevos instrumentos que armonicen con esas concepciones y de sentirse como evaluador. Alguien que realiza catas o sondeos regulares del estado del aprendizaje para optimizarlo, teniendo claro su objetivo y sintiéndolo como una necesidad y no, o no solo, alguien que examina, juzga y dictamina sobre el futuro del aprendiz.

Para lograr una reconstrucción efectiva de la posición evaluadora de los docentes debemos conocer cuáles son sus representaciones más frecuentes ante la evaluación para, posteriormente, tratar de introducir tensiones e incidentes que promuevan cambios en esas posiciones, siguiendo el modelo de propuestas de formación existentes centradas en este proceso de reconstrucción de posiciones a través del análisis de incidentes críticos (Monereo, Weise et al., 2013; Weise y Sánchez-Busqués, 2013).

Por ello, en este estudio analizaremos, en primer lugar, qué posicionamientos están en consonancia con una evaluación más tradicional, vinculada a la “cultura del test” antes mencionada, y cuales se corresponden con un enfoque de evaluación auténtica, sintonizada con el modelo pedagógico actual basado en el desarrollo de las competencias y la evaluación orientada al aprendizaje (Sanmartí, 2007; Monereo, 2019). En segundo lugar, exploraremos si este posicionamiento guarda relación con los instrumentos diseñados durante la formación sobre evaluación auténtica y si se correlaciona con la voluntad de seguir utilizándolos. Es decir, queremos explorar si el cambio en el posicionamiento docente sobre evaluación viene precedido por un cambio en las prácticas o le es consustancial. Este objetivo

nos llevará a justificar la posibilidad de transitar de unas a otras posiciones identitarias mediante un sistema basado en dos dispositivos fundamentales:

- en un proceso de autenticación de las pruebas de evaluación que habitualmente emplean los profesores; a través de la Guía GAPPISA (*Guía para el Análisis de pruebas desde la perspectiva de Pisa*).
- mediante la toma de consciencia de los incidentes que se han producido durante el proceso como fruto de la tensión entre las actividades de enseñanza, las pruebas de evaluación modificadas y los resultados obtenidos por los alumnos.

Seguidamente analizaremos con mayor profundidad ambos dispositivos de cambio.

3.1.1. Dispositivos para el cambio de posicionamiento en evaluación

Como hemos comentado anteriormente, cambiar la propia posición sobre la evaluación implica modificar el significado y el sentido que ésta tiene para el docente, es decir, sus concepciones. Simultáneamente, requiere que el docente adopte nuevos métodos de evaluación acordes con esas concepciones. Y, por último, también paralelamente, estos cambios le exponen a acompañar esas ideas y decisiones con emociones que generen expectativas positivas sobre sus habilidades y se enfoquen a afrontar asertivamente momentos de cambio e incertidumbre, inherente a una evaluación centrada en los procesos y no solo en los conocimientos.

Partimos pues, de una conceptualización de la identidad como el conjunto de representaciones que declara y pone en acción el docente sobre las creencias respecto a lo que significa su rol profesional y cómo debe actuar y los sentimientos que emergen en esta relación, teniendo este conjunto componentes estables en el tiempo y componentes dinámicos según el contexto (Day, Kington, Stobart y Sammons, 2006; Monereo y Pozo, 2011; Weise y Sánchez-Busqués, 2013). Y el primer dispositivo de cambio consiste precisamente en tomar consciencia de estas dimensiones de la identidad, para poder posicionarse y enfocarlos a una práctica de evaluación realmente auténtica. Una manera de conseguirlo consiste en

provocar un cambio a través de la formación práctica y reflexiva de los docentes (Liesa et al., 2018). Un ejemplo de ello es el proyecto de formación GAPPISA, expuesto en el capítulo anterior, donde el docente transforma sus pruebas de evaluación, las aplica y reflexiona sobre su impacto real en el aula (Mejías y Monereo, 2018).

Las características de la guía GAPPISA, ponen en cuestión una concepción tradicional de la evaluación y sus estrategias, y a través de indicadores de cambio la mueve hacia un enfoque competencial, formativo e inclusivo, es decir, de evaluación auténtica. Este dispositivo permite reflexionar sobre qué significa evaluar y qué debemos evaluar, cómo, quién y en qué momento debe hacerse y cuáles son las expectativas y emociones reguladoras en el proceso. A continuación, describimos estas dimensiones con más detalle, especificando algunas de las representaciones más comunes que tienen los docentes, con el fin de conceptualizar los dos modelos, el enfoque de evaluación tradicional y el enfoque de evaluación auténtica.

A nivel de concepciones, se trata de identificar si se transita por un polo social en el que priman las funciones evaluativas centradas en acreditar los aprendizajes y rendir cuentas sobre la enseñanza impartida y, por lo tanto, una evaluación básicamente de conocimientos, o por un polo pedagógico en el que priman las funciones centradas en la evaluación en y de los procesos de enseñanza y aprendizaje, para que los docentes mejoren su enseñanza y el alumnado su aprendizaje participando en la evaluación y, por lo tanto, una evaluación principalmente de competencias (Barnes et al., 2017; Brown y Remesal, 2012; Coll y Remesal, 2009; Levy-Vered y Nasser-Abu, 2018; Lumsden y Michaelides, 2011; Sanmartí, 2010b; Segers y Tillema, 2011).

A nivel de estrategias, consiste en rechazar la idea de una evaluación única al final de la Unidad Didáctica impartida, en la que se apliquen instrumentos tipo test o exámenes de respuesta cerrada, bajo el pretexto de que la información recogida sea lo más fiable posible. También dejar de ver el docente como único experto en los conocimientos que determinan los diferentes niveles a alcanzar. Por el contrario, sí se trata de sostener una evaluación que se realiza antes, durante y después de

la Unidad Didáctica o proyecto, aplicando instrumentos diseñados para analizar las evidencias de aprendizaje e identificar las dificultades y superarlas, sosteniendo el lema de que el error es una oportunidad de aprendizaje. En este enfoque, los docentes evalúan a los alumnos y se evalúan a sí mismos y, a la vez, el alumnado evalúa a sus compañeros, se evalúa a sí mismo y, eventualmente, incluso al docente. Para ello los criterios de evaluación están abiertamente consensuados y compartidos previamente. En definitiva, una evaluación integrada en los procesos de enseñanza y aprendizaje y que promueva la competencia de aprender a aprender y respete los perfiles de aprendizaje (Black, 2015; Coll, Mauri y Rochera, 2012; Morodo, 2017; Swaffield, 2011; Tierney, 2006).

A nivel de sentimientos, como el alumnado debe alcanzar unos niveles consensuados para demostrar que la enseñanza ha sido de calidad, el docente normalmente se siente con el deber de que la evaluación sea justa en tanto que es para todos igual y nadie recibe ayudas específicas para poder garantizar la objetividad de esta. Sin embargo, el docente que escoge una evaluación auténtica siente que la evaluación es una actividad primordial durante los procesos de enseñanza y aprendizaje y que, por lo tanto, debe mantener una actitud de confianza en sus alumnos y de voluntad de cambio y mejora constantes, desde una visión de equidad y no exclusión (Kaur et al., 2015; Kaur et al., 2017).

A partir de tomar conciencia, como señalábamos anteriormente, de las representaciones que conforman sus concepciones sobre evaluación, y de ponerlas en práctica modificando y aplicando un instrumento de evaluación, el docente pone al descubierto su posicionamiento sobre evaluación. De esta manera, podemos identificar cuál es el perfil de evaluación de cada docente (Coombs et al., 2018), y si tiene a un enfoque de evaluación tradicional o a un enfoque de evaluación auténtica, aunque sabemos que todas las características que los definen son complementarias entre ellas y forman un continuo (Postareff, Virtanen, Katajavuori, Lidblom-Ylänne, 2012).

Por otra parte, debemos tener presente que esa posición evaluativa que mantiene todo docente está en continuo diálogo con otras posiciones profesionales como la de instructor, tutor, gestor de normas de conducta, integrante de un equipo, etc.

Desde el modelo teórico del que partimos, la teoría del *self* dialógico postulada por Hubert Hermans (2001, 2009, 2014), la identidad está conformada por un conjunto de posiciones en diálogo, cuya interacción no siempre resulta armoniosa y complementaria. Un profesor puede presentar una posición muy autoritaria, auto-centrada y objetivista como evaluador y, sin embargo, en calidad de instructor posicionarse como un mediador, centrado en las ideas de los alumnos, tolerante y dialogante en clase, y partidario de metodologías más inductivas y activas. Lógicamente esas divergencias entre posiciones suelen desembocar en tensiones e incidentes, es decir, en situaciones emocionalmente desequilibrantes que pueden llegar a poner en crisis al docente y llevarlo incluso a tomar decisiones extremas - por ejemplo, abandonar la profesión- (Bilbao y Monereo, 2011).

La presencia de esos incidentes, en especial cuando son críticos y fuerzan a un cambio a quien los padece, han demostrado una gran eficacia en situaciones formativas intensivas cuyo objetivo es promover un cambio radical en la identidad profesional de los participantes (Monereo, Weise y Álvarez, 2013). Se ha demostrado que la experiencia de superar un incidente crítico es una oportunidad para adoptar una nueva posición estratégica que impacta en emociones y prácticas (Monereo, 2017; Weise y Sánchez-Busqués, 2013). En este trabajo también exploramos los incidentes que declaran experimentar los docentes sobre evaluación para verificar su eficacia como dispositivo de cambio.

Sostenemos que la revisión y explicación de los incidentes que los docentes han tenido como fruto de la aplicación de la prueba auténtica y la discrepancia y problemáticas que esto ha generado en las actividades de enseñanza y los resultados de los alumnos, permite tomar consciencia sobre la manera de pensar, hacer y sentir la evaluación que tiene el docente y, por lo tanto, redefinirla a través de modificaciones en el discurso y en la práctica (Álvarez-Bernárdez y Monereo, 2016; Badia y Chumpitaz-Campos, 2018; Monereo, Castelló et al., 2009; Monereo y Morcillo, 2011).

En síntesis, los objetivos que guían esta investigación son:

1. Caracterizar el posicionamiento sobre evaluación de los docentes de enseñanza de primaria y secundaria, según se sitúen en un enfoque de evaluación tradicional o en un enfoque de evaluación auténtica.
2. Explorar la relación entre los posicionamientos docentes sobre evaluación y la calidad y la permanencia de las pruebas diseñadas durante la formación.

3.2. Método del Estudio.

En el diseño de esta investigación, considerando la muestra con la que contábamos y el hecho de no poder acceder a ella físicamente (la formación sobre evaluación auténtica se realizó durante tres cursos escolares y de forma virtual puesto que participaban escuelas distribuidas por toda España), optamos por utilizar la aplicación de un cuestionario online y no aleatorio. El cuestionario, junto a la entrevista, es uno de los instrumentos más utilizados cuando se trata de analizar la voz de los sujetos y todo lo que declaran. El hecho de realizarse por internet responde más a una necesidad debido a la dificultad de contactar con la muestra y la disponibilidad de esta -los docentes estaban distribuidos por toda España y había pasado mínimo un año desde su participación en el curso-, que, a una estrategia de investigación, con las limitaciones que ello comporta.

3.2.1. Contexto y procedimiento.

La investigación se contextualiza en la implementación de la formación sobre evaluación auténtica a docentes de educación primaria y secundaria *Proyecto GAPPISA*, explicada en el capítulo anterior. Recordamos que el principal objetivo fue dotar de estrategias y recursos a los docentes para que pudieran modificar y diseñar una prueba de evaluación que cumpliera los criterios característicos de la autenticidad, fundamentales para la evaluación de competencias. Por lo tanto, los resultados con los que se contó fueron pruebas de evaluación auténticas diseñadas por los docentes y evaluadas por los asesores.

Por otro lado, lo que nos interesa destacar de nuevo en este apartado, es que la formación se llevó a cabo durante tres años consecutivos en varias escuelas del territorio español que formaban parte de dos grupos diferenciados de redes de

centros educativos. Reunimos todos los datos de estos participantes sin diferenciar red de escuelas y año de formación porque consideramos que no eran variables determinantes puesto que la formación siempre tenía la misma estructura. Contactamos con ellos para solicitar su participación en el estudio de forma virtual; repetimos la demanda en dos ocasiones, distribuidas temporalmente en diferentes trimestres.

3.2.2. Participantes.

La muestra sobre la que se basa este artículo la conforman 20 docentes. Del total de respuestas recibidas, seleccionamos a los participantes según las siguientes características: que hubieran realizado todas las actividades de la formación sobre evaluación auténtica para verificar que los resultados fueran fiables; que fueran maestros de educación primaria y secundaria por ser el grupo más representativo; y que procedieran de diferentes ediciones (desde el curso 2012 al curso 2015) para mostrar la diversidad en la permanencia de los cambios en caso de que la hubiera.

3.3.3. Instrumentos de recogida y análisis de datos.

Los instrumentos de recogida de datos utilizados fueron los siguientes:

- *Rúbrica de evaluación de pruebas auténticas* (ver Anexo 2): instrumento utilizado durante la formación en la revisión de pruebas por parte del equipo asesor. En ella se describen las variables de competencialidad, carácter formativo e inclusividad, definidas en capítulos anteriores, las cuales conforman el modelo de pruebas de evaluación auténticas. Cada variable se detalla en cuatro niveles distintos de consecución según los cambios realizados. De esta manera, con el instrumento se puede determinar la puntuación de cada variable y del total de la prueba según el nivel de logro alcanzado.
- *Cuestionario “Mi posicionamiento sobre evaluación”* (ver Anexo 5): instrumento dividido en tres apartados; el primero, sobre la participación durante la formación, el segundo, relacionado con datos educativos sobre el sujeto y, el tercero, centrado en la evaluación en las dimensiones de concepciones, estrategias, sentimientos e incidentes. Está compuesto por

un total de 22 preguntas (21 cerradas y 1 abierta), que reflejan las principales variables sobre el posicionamiento en evaluación identificadas en la literatura.

En la siguiente tabla se muestra una relación entre las dimensiones de análisis, las variables y los indicadores utilizados (ver Tabla 27). Respecto al análisis de datos, para obtener los datos de la dimensión resultados de formación se realizó un análisis descriptivo de los mismos. Para el análisis del contenido del cuestionario se utilizó un instrumento diseñado específicamente para ello a partir de la literatura sobre identidad docente y evaluación tradicional y auténtica (ver Anexo 6). Este instrumento permite describir cuál es la tendencia del posicionamiento docente sobre evaluación, realizando un análisis descriptivo de esta.

Tabla 25

Relación entre las dimensiones, las variables y los indicadores de análisis del estudio

Dimensiones	Variables	Indicadores
Resultados formación	Nivel de la prueba auténtica	- Puntuación versión final
	Permanencia	- Declara aplicar su prueba GAPPISA - Declara diseñar otras siguiendo el modelo GAPPISA.
Concepciones	Finalidad de la evaluación	- Acreditación de niveles alcanzados (polo social) - Rendición de cuentas sobre resultados (polo social) - Recoger indicadores para mejorar la enseñanza (polo pedagógico) - Ofrecer ayudas para que los alumnos se corrijan y mejoren (polo pedagógico)
	Objeto de evaluación	- Conocimientos - Estrategias de aprendizaje - Competencias - Aprendizajes y actuación docente
Estrategias	Enfoque a nivel de centro	- Según las exigencias externas - Según una programación prediseñada por otros agentes - Cada docente a su criterio - Consenso del equipo docente
	Instrumentos de evaluación	- Exámenes de ejercicios y problemas sobre conocimientos - Exámenes tipo test o preguntas de diferente tipología - Evaluación basada en la resolución activa de problemas realistas y competencias - Portafolio con registro de todas las evidencias

	Momento de la evaluación	<ul style="list-style-type: none"> - Después de la UD o proyecto - A lo largo de la UD o proyecto - Antes de las UD o proyecto - Durante las actividades para ajustar la ayuda
	Agente que evalúa	<ul style="list-style-type: none"> - Instancias externas - Varios: docentes, alumnado, otros agentes externos - Únicamente el docente - Autoevaluación y coevaluación
	Características de las pruebas	<ul style="list-style-type: none"> - Interdisciplinariedad - Tareas y/o preguntas auténticas y contextualizadas - Graduación en la exigencia cognitiva creciente - Formato de pregunta heterogéneo - Condiciones de resolución auténticas - Ajuste a la diversidad: actividades optativas - Preguntas sobre el proceso de resolución - Tareas de autoevaluación y/o coevaluación - Consenso previo de los criterios de evaluación - Comunicación y correcciones como oportunidad de aprendizaje
	Satisfacción	<ul style="list-style-type: none"> - Nivel a través de una puntuación del 1 al 10.
Sentimientos	Sentimiento imperante	<ul style="list-style-type: none"> - Con motivación y ganas de compartir - Con presión - Con ganas de experimentar y mejorar - Con un sentimiento de impartir justicia
	Preocupaciones principales	<ul style="list-style-type: none"> - Valorar las competencias adquiridas - Adjudicar una calificación justa - Recoger indicadores para ajustar la ayuda educativa - Aportar evidencias útiles para que los alumnos puedan mejorar
Incidentes	Tipología de incidentes	<ul style="list-style-type: none"> - Incidentes referentes a reflexiones sobre evaluación. - Incidentes que suceden durante la aplicación de la evaluación. - Incidentes que suceden una vez aplicada la evaluación.

3.3. Resultados

A continuación, describimos los resultados en dos grandes grupos: primeramente, caracterizaremos el posicionamiento de los docentes con relación a la evaluación tanto para un enfoque más tradicional como para uno más auténtico; segundo, mostraremos la correspondencia ente los posicionamientos declarados y los resultados de la formación con relación a la permanencia y calidad de las pruebas.

Antes de empezar, es preciso señalar que los resultados que vamos a exponer son descriptivos dado parten del propósito del estudio es caracterizar a la muestra seleccionada del estudio anterior sobre los resultados del proyecto de formación GAPPISA, y deducir las implicaciones prácticas para mejorar. También, pretenden ser una primera aproximación al estudio de las posiciones sobre evaluación desde el paradigma de la identidad docente, que permita diseñar estudios posteriores válidos para una muestra más amplia en la que sí se puedan realizar generalizaciones.

Además, dado que no ha habido un patrón característico que correlacione las dimensiones en los diferentes casos, en lugar de describir todos los perfiles, hemos analizado los datos a la luz de la totalidad de la muestra, de manera que hablamos del grupo y no de los casos. Por otro lado, sin embargo, sí nos ha parecido adecuado señalar algunos casos en un apartado específico, por ser los más dispares.

3.3.1. Posicionamiento de los docentes sobre evaluación

En este primer punto, exponemos los resultados obtenidos en cuatro niveles de análisis, siempre describiendo cuál es la tendencia, si hacia representaciones definitorias de una evaluación más tradicional o de una evaluación más auténtica. En un primer nivel, expondremos la tendencia a un enfoque tradicional o auténtico de la evaluación que tienen el conjunto de los participantes; en un segundo nivel, la tendencia de éstos en cada una de las dimensiones -concepciones, estrategias y sentimientos-; en un tercer nivel, describiremos las variables que definen cada una de las dimensiones; y, en un cuarto nivel, las características que predominan en cada una de las variables y las relaciones entre éstas.

Primero, definimos una tendencia hacia un enfoque tradicional o auténtico agrupando las representaciones sobre evaluación que declaran tener los docentes y que hacen referencia a dicho enfoque. En el caso que las representaciones totales no sean representativas en su conjunto de un enfoque, los definimos como sin tendencia clara. A la luz del análisis realizado, podemos decir que la mayoría de los docentes declaran tener en su conjunto una posición sobre evaluación que tiende

a un enfoque auténtico (ver Tabla 28). Por el contrario, algunos docentes siguen anclados en un enfoque tradicional. Por último, un menor porcentaje mantiene un posicionamiento que no tiene una tendencia clara.

Tabla 26

Posicionamiento sobre evaluación de los docentes según su tendencia

	Sujetos	Frecuencia
Enfoque tradicional	3	15%
Enfoque auténtico	15	75%
Sin tendencia clara	2	10%
Total	20	100%

Es necesario, por tanto, analizar las dimensiones de concepciones, estrategias y sentimientos para descubrir con un mayor detalle qué sucede (ver Figura 7). De esta manera, encontramos que, con respecto a las concepciones, los docentes se sitúan tanto en un enfoque tradicional como en uno auténtico. Sin embargo, a nivel de estrategias de evaluación, el doble de los docentes declara que realizan una evaluación auténtica. Por último, con relación a los sentimientos, esta tendencia aumenta notablemente, de forma que más del doble de los docentes se declara en una posición más auténtica.

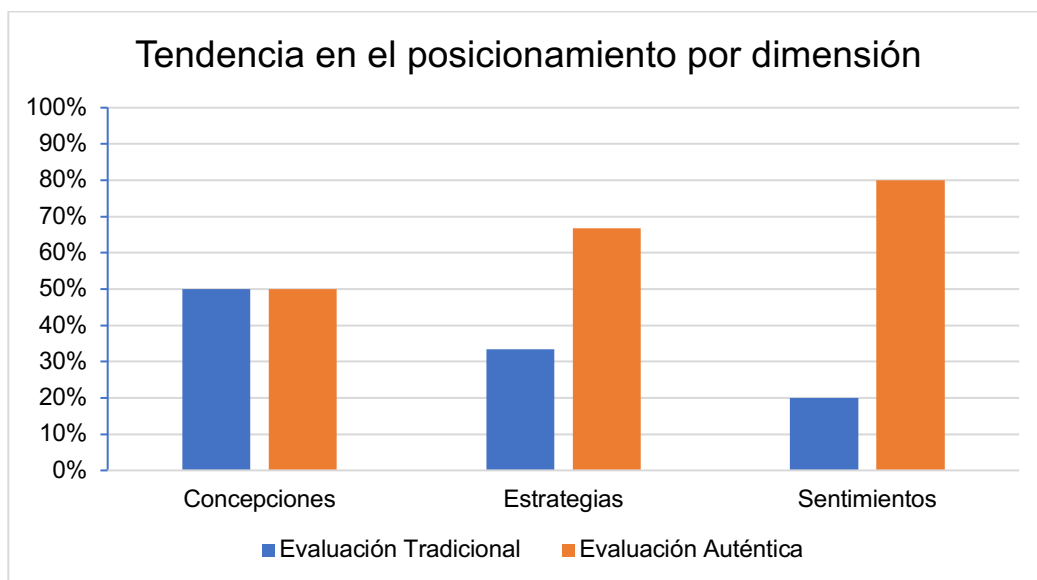


Figura 7. Posicionamiento sobre evaluación de los docentes por dimensión

Realizamos un paso más y analizamos las variables que definen cada dimensión para caracterizar en profundidad dicho posicionamiento (ver Figura 8); observamos que a nivel de las variables que definen cada dimensión, en todas menos en una hay una tendencia hacia representaciones coincidentes con una evaluación auténtica. La excepción se sitúa en la dimensión de concepciones y hace referencia al objeto de la evaluación, en la que la mayor parte de los docentes declararon que eran los conocimientos y el nivel de competencias.

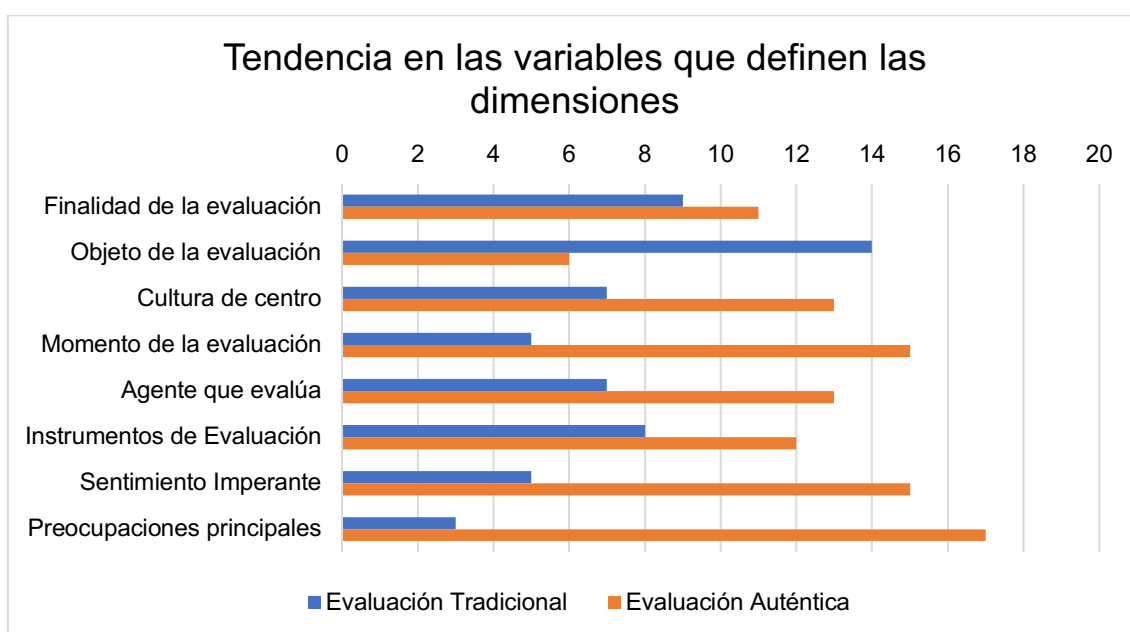


Figura 8. Tipo de evaluación según las categorías que conforman las dimensiones del posicionamiento docente sobre evaluación.

Es preciso pues, examinar los resultados para cada una de las variables, señalando cuáles son los indicadores que han obtenido mayor puntuación. Queremos describir estas variables para poder luego analizar cómo influyen en la consistencia interna de cada dimensión, es decir, para conocer qué aspectos distorsionan un posicionamiento y discutir posteriormente las repercusiones.

En primer lugar, si prestamos atención a las características que definen la dimensión de concepciones -la finalidad, el objeto y la cultura de centro-, dos de ellas generan una contradicción interna dentro de la dimensión: la finalidad de la evaluación (ver Figura 9) con el objeto de esta (ver Figura 10).

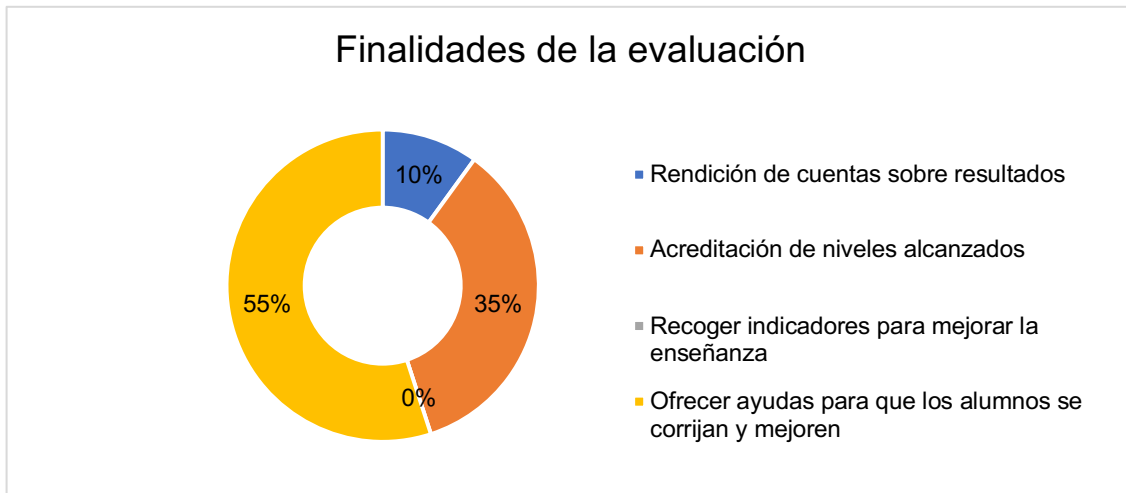


Figura 9. Tendencia en la variable finalidad de la evaluación.

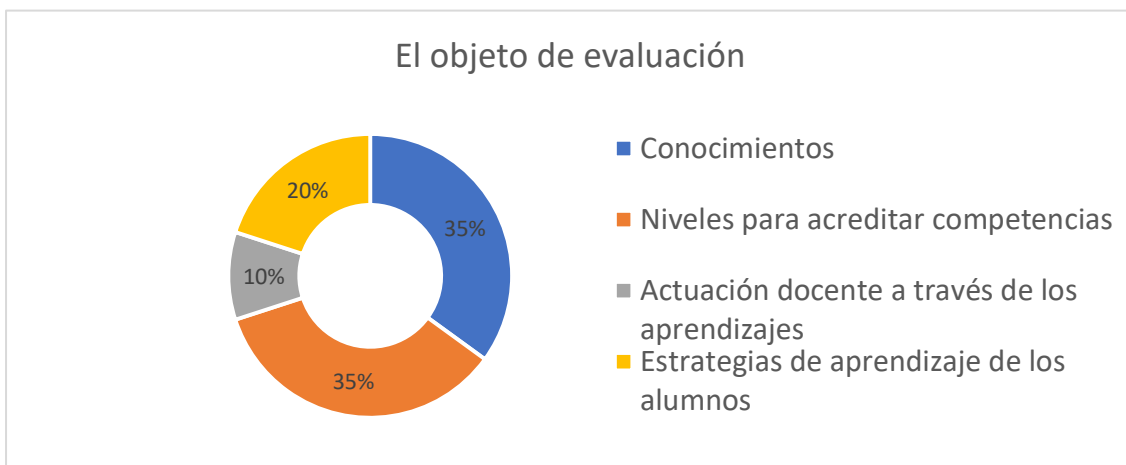


Figura 10. Tendencia en la variable objeto de la evaluación.

Mientras un 55% de la muestra declara que la finalidad de la evaluación es “ofrecer ayudas para que los alumnos se corrijan y mejoren”, solo el 20% afirma que el objeto de la evaluación es trabajar sobre las “estrategias de aprendizaje de los alumnos y sus procesos”, en contraposición a un 35% que asevera que el objeto de la evaluación son los conocimientos y otro 35% los niveles para acreditar competencias.

En segundo lugar, si analizamos las características que definen la dimensión de estrategias -quiénes evalúan, cuándo y con qué instrumentos-, comprobamos que también hay cierta discrepancia entre las características con mayor consenso. En este caso no hay una coherencia entre los instrumentos que los docentes declaran que aplican (ver Figura 12) y el agente implicado en su uso (ver Figura 11).

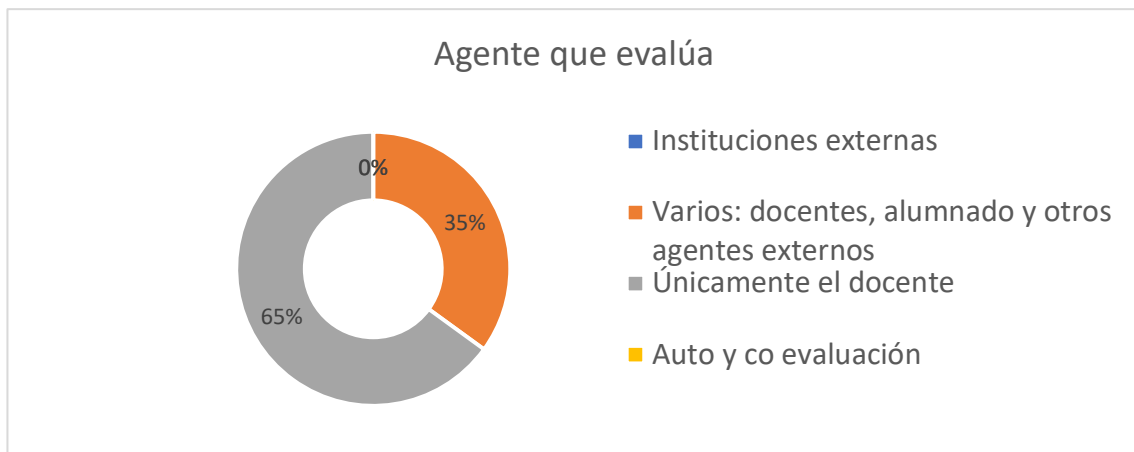


Figura 11. Tendencia en la variable agente que evalúa

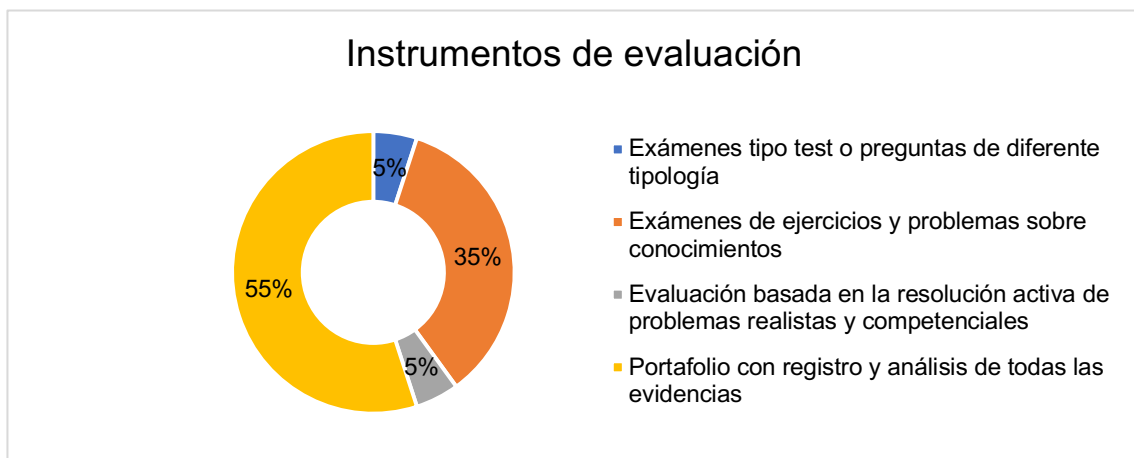


Figura 12. Tendencia en la variable instrumentos de evaluación

El instrumento más utilizado es el Portafolio, el cual requiere que principalmente sea el alumno quien recoja, seleccione, analice y valore sus evidencias de aprendizaje, en definitiva, que sea el alumno quién evalúe. Sin embargo, los docentes declaran mayoritariamente que es únicamente el profesor quien evalúa y no los alumnos. Respecto a los momentos en los que se realiza la evaluación sí

podríamos decir que se evalúa a lo largo de la misma, lo cual favorece un enfoque auténtico.

En tercer lugar, si detallamos la dimensión relativa a los sentimientos, el primer resultado se orienta a la satisfacción general respecto a su evaluación con una declaración mayoritariamente positiva. La segunda cuestión, que refuerza este aspecto, es que los docentes se sienten motivados respecto a la tarea de evaluar y, además, desean sostener una visión compartida de la evaluación (ver Figura 13).

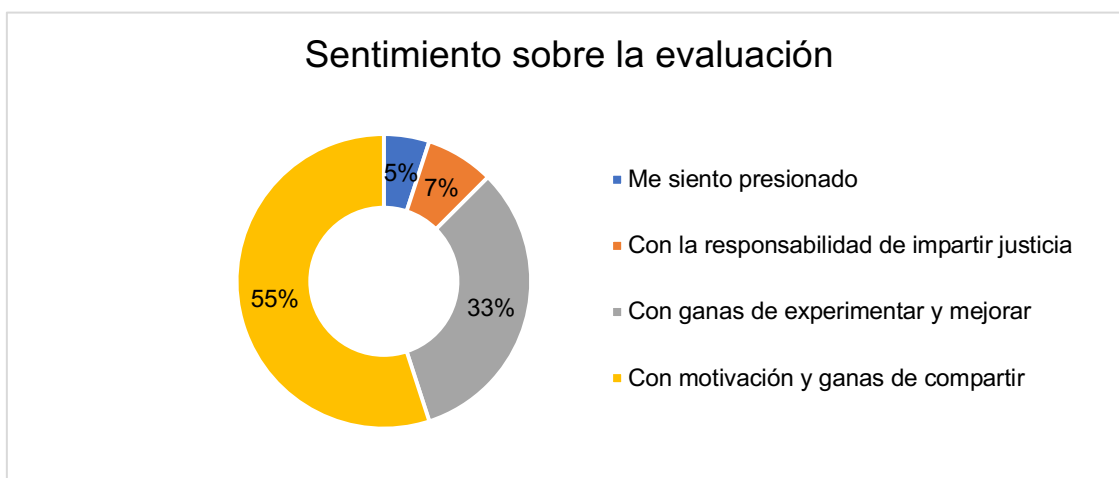


Figura 13. Tendencia en la variable sentimientos respecto a la evaluación

Centrándonos en las preocupaciones que declaran, los profesores sienten que quieren mejorar y hacerlo en la dirección de una evaluación formativa que aporte herramientas para ayudar a los alumnos. Esto se relaciona con el hecho de que, en las dos variables anteriores, concepciones y estrategias, se declaren contradicciones relacionadas precisamente con empoderar a los alumnos para que sean protagonistas de su evaluación y se transforme ésta en una experiencia de aprendizaje.

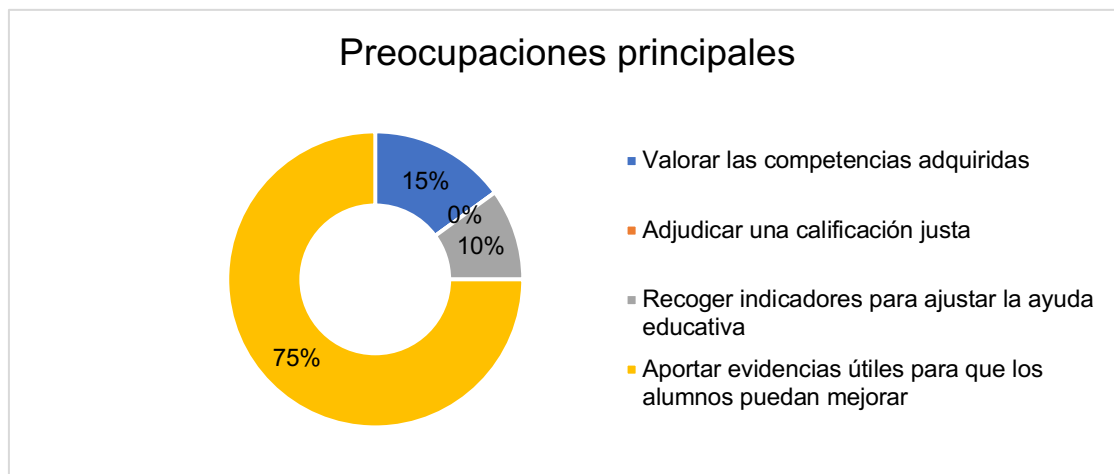


Figura 14. Tendencia en la variable preocupaciones sobre evaluación

En conclusión, los resultados nos muestran que el posicionamiento docente sobre evaluación tiende, mayoritariamente, hacia un enfoque de evaluación auténtica con sentimientos de motivación e interés por compartir la experiencia, aunque los docentes aún declaran tener creencias y actividades de evaluación propias más de un enfoque tradicional, tensiones que se corroboran con las preocupaciones que detallan por conseguir, de verdad, una evaluación auténtica. También nos muestran que no todas las variables que forman las categorías están alineadas hacia una misma posición sobre evaluación y, por lo tanto, las dimensiones siguen en construcción.

3.3.2. Incidentes sobre evaluación

El siguiente aspecto por describir son los incidentes que los profesores explican que han vivido a raíz de aplicar la prueba auténtica diseñada y en su día a día como evaluadores innovadores. La primera observación es que algunos describieron sus incidentes en positivo, es decir, describieron experiencias que les habían impactado de forma positiva. Éstas estaban relacionadas con un aumento de la motivación por parte de los alumnos, con el aprendizaje profundo que estos consiguieron y, de forma más específica, con la emoción de un alumno al verse reflejado en una situación realista de la prueba auténtica.

Presentando ahora los resultados sobre los incidentes críticos relacionados con la evaluación que impactaron negativamente en los docentes, los clasificamos en tres grupos: los que hacen referencia a reflexiones sobre evaluación, los que mencionan situaciones que suceden durante la aplicación y los que se refieren a experiencias ocurridas una vez aplicada la prueba (ver Tabla 29).

Tabla 27

Tipos de incidentes críticos con relación a la evaluación

Incidentes referentes a reflexiones sobre evaluación	Incidentes que suceden durante la aplicación de la evaluación	Incidentes que suceden una vez aplicada la evaluación
Dificultad en cumplir el currículum y en ajustarse a la ley educativa.	Dificultad en aplicar tareas de coevaluación (alumnos no imparciales y alumnos sinceros).	Resistencias dentro del equipo docente sobre la nueva evaluación.
Dificultad en evaluar de acuerdo con la metodología de aula.	Dificultad en el uso de instrumentos (rúbrica) para autoevaluación.	Quejas de los alumnos por estar en desacuerdo con su nota.
Dificultad en seleccionar resultados acordes a los procesos de aprendizaje.	Falta de estrategias para la resolución tanto de pruebas sencillas como de reflexión.	Alumnos con muy buenas notas que a en la prueba competencial bajan su puntuación.

Los docentes se ven afectados por la distancia entre la prueba auténtica y las exigencias externas o las prácticas educativas: por un lado, a causa de tener que cumplir con el currículum marcado o los estándares de aprendizaje que marca la ley; por otro lado, debido a la aplicación de una metodología no acorde a la evaluación. En palabras de los docentes: *“Insatisfacción por la incorrecta forma de evaluar que acompaña al método seguido en las clases”* (S1, E1); *“No recoger resultados acordes con los procesos de los alumnos”* (S7, E1).

Además, y relacionado con las estrategias de evaluación, varios docentes manifiestan haber vivido sucesos que no esperaban en el momento de llevar a cabo tareas de autoevaluación y de coevaluación como, por ejemplo, alumnos que valoraban los trabajos de los compañeros según la amistad que tenían y no de forma imparcial. O el hecho de que los alumnos no utilizaran la rúbrica para autoevaluarse y que, previamente, el docente la hubiese compartido con las expectativas puestas en que lo hicieran.

También ha resultado un incidente la falta de habilidades y conocimientos para realizar actividades sencillas y de reflexión durante la realización de la prueba (como, por ejemplo, dificultad para leer los enunciados e identificar la tarea, pocas habilidades en expresión escrita, etc.).

Y, en último lugar, en referencia al impacto que tiene la prueba en docentes, familias y en los propios alumnos, ha resultado que las resistencias dentro del claustro docente suponen claramente una dificultad desestabilizadora, en palabras de un docente: “...encontrarme con otros docentes que no aprueban esta nueva forma de trabajar con los alumnos.” (S2, E1), hecho que no permite la construcción de una cultura de centro sobre la evaluación. Los incidentes con los alumnos están más relacionados con los resultados, en tanto que ha habido cambios en las calificaciones, por ejemplo, que aquellos que obtenían muy buenas puntuaciones hayan bajado la nota a raíz de la aplicación de la prueba competencial.

3.3.3. Permanencia en el diseño y aplicación de la prueba auténtica y competencial.

Finalmente, vamos a analizar la permanencia según las variables definidas. Para comenzar, veremos si declaran aplicar la prueba auténtica y si confirman que diseñan pruebas desde un enfoque auténtico. Seguidamente, si las pruebas o actividades de evaluación que aplican actualmente mantienen de forma mayoritaria las características del modelo GAPPISA, según su descripción.

Gran parte de los docentes, un 80% de la muestra, declaran aplicar la prueba diseñada ajustándola al grupo de alumnos con el que trabajan. Sin embargo, muchos más docentes declaran seguir diseñando pruebas de evaluación en base a los criterios de la Guía GAPPISA (ver Figura 15).

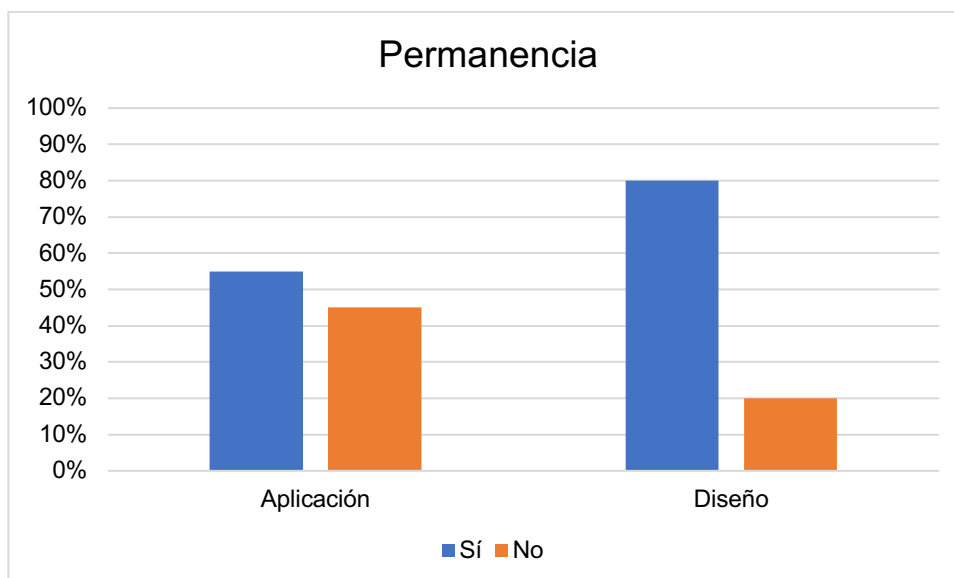


Figura 15. Permanencia sobre el diseño y la aplicación de pruebas que cumplen las características de la Guía GAPPISA

Seguidamente, analizamos como caracterizan los profesores las pruebas y actividades de evaluación que de forma mayoritaria aplican en el aula. Les pedimos que seleccionasen aquellas que eran representativas de su forma de evaluar para relacionar cuántas coincidían con el modelo GAPPISA, más allá de que aplicasen o no la prueba. El resultado es que más de la mitad de la muestra no cumple la mitad de los criterios característicos del modelo GAPPISA, es decir, no los aplica mayoritariamente en la gran variedad de actividades de evaluación que realizan.

Es interesante pues conocer cuáles son las características del modelo GAPPISA que los docentes han podido integrar en sus prácticas de evaluación de forma representativa: compartir los criterios de evaluación, proponer tareas de autoevaluación y coevaluación, heterogeneidad en el formato de enunciado y respuestas y una graduación de la exigencia cognitiva. A estas características (ver Figura 16), con un nivel menor de consenso, les siguen una comunicación de los resultados y correcciones entendida como una nueva oportunidad de aprendizaje y diseñar tareas y/o preguntas realistas, relevantes y socializadoras que se refieran a una situación inicial a resolver bien contextualizada.

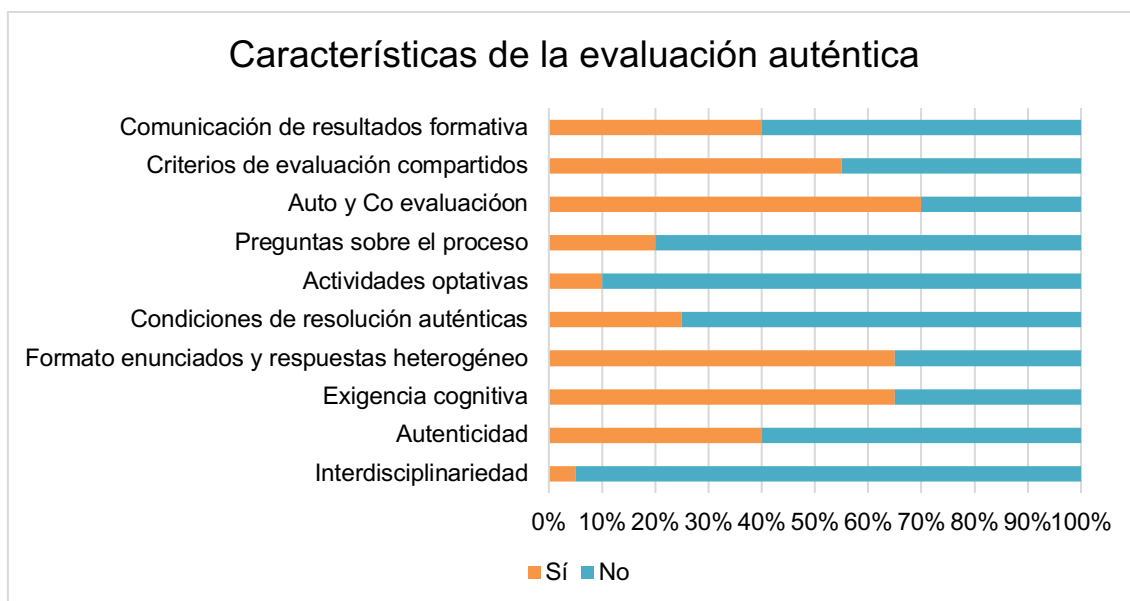


Figura 16. Características que cumplen las pruebas y actividades de evaluación de los profesores según el modelo GAPPISA

Por último, hemos relacionado cada uno de los perfiles, tanto los que tienen un posicionamiento más cercano a una evaluación tradicional, como los que se posicionan cerca de una evaluación auténtica, con las variables de permanencia, nota final de la prueba de evaluación y años desde la aplicación de la prueba. En ninguno de los casos hemos encontrado un patrón que se repita, es decir, aunque el docente en su conjunto y de forma declarativa tienda a una evaluación tradicional, esto no significa que no continúe aplicando o diseñando pruebas competenciales, según haya pasado más o menos tiempo desde que realizó la formación y/o esté más o menos satisfecho por su evaluación.

Tampoco hemos encontrado resultados significativos entre las variables mencionadas y los años de experiencia docente, la materia o el nivel educativo donde se imparte docencia. Por último, si relacionamos cuántos años hace que se diseñó la prueba y la continuidad en la aplicación o diseño de nuevas pruebas, tampoco encontramos resultados claros.

3.4. Discusión de los resultados.

En este apartado vamos a discutir los resultados a la luz de la literatura sobre la temática para dar respuesta a los objetivos planteados y proponer nuevas líneas

de investigación. Primeramente, discutiremos la caracterización del posicionamiento sobre evaluación describiéndolo en relación con la evaluación tradicional y la evaluación auténtica. En segundo lugar, exploraremos el peso que tienen los incidentes críticos en la construcción de las posiciones evaluativas. En tercer lugar, reflexionaremos sobre la relación entre las posiciones identificadas y la calidad de las pruebas de evaluación que realizaron y la permanencia de su uso. Y, para acabar, expondremos las implicaciones a nivel práctico que de este análisis de resultados pueden extraerse.

Respecto al primer objetivo, la caracterización del posicionamiento docente sobre evaluación atendiendo a las dimensiones de esta y las tensiones internas entre las variables de estas dimensiones, encontramos variedad de respuestas. El resultado principal es que, de forma mayoritaria, el posicionamiento docente sobre evaluación que declaran los profesores de enseñanza primaria y secundaria que participaron en la formación GAPPISA hace un máximo de tres años, tiende hacia posiciones coincidentes con una evaluación auténtica, tanto en lo que se refiere al conjunto de concepciones, estrategias y sentimientos como a cada una de las variables que componen estas dimensiones de la identidad. A continuación, vamos a destacar y discutir las principales características de dicho posicionamiento. Cabe recordar que nos centramos en la identidad declarada en relación con la evaluación, es decir, el posicionamiento docente sobre evaluación, analizando sus concepciones, estrategias y sentimientos.

A nivel de concepciones, entendidas como el conjunto de teorías, más o menos explícitas, en torno a la evaluación, hemos analizado aquello que piensan los profesores sobre cuáles son las finalidades de la evaluación, por qué evalúa, qué debe evaluar principalmente y si debería haber un consenso a nivel de centro respecto a la evaluación. Los resultados nos muestran que solo la mitad de los profesores se declaran favorables a una evaluación auténtica, aunque de forma no consistente, en tanto que sostienen creencias diferentes y discrepantes entre ellas, como ya encontraron otros estudios (Remesal, 2011).

Si analizamos dichas tensiones internas, constatamos que, aunque los profesores tienen claro que el propósito de la evaluación es que los alumnos se corrijan y

mejoren, es decir, una evaluación para el aprendizaje, pocos sostienen que deberían evaluar las estrategias y los procesos de los alumnos. Es decir, los docentes sitúan al alumno como protagonista cuando piensan para qué sirve la evaluación, pero no facilitan que así sea cuando deciden cómo va a ser el proceso de evaluación, en el que eligen instrumentos más relacionados con medir y acreditar los aprendizajes. A la luz de estas tensiones, estamos de acuerdo con Coll y Remesal (2009) cuando señalan en su estudio que las tendencias dentro de estas concepciones pueden estar condicionadas por las tensiones entre las diferentes actividades, pedagógicas y sociales, a las que se ven obligados los docentes.

Estas resistencias al cambio de paradigma en evaluación son comprensibles si tenemos en cuenta que el discurso sobre evaluación auténtica es todavía reciente e incluye redefinir muchos de los principios evaluativos sostenidos hasta el momento (Ahumada, 2005), siendo los docentes conscientes de las resistencias al cambio (Syrene, Veville y Chapman, 2009). Por otro lado, debemos recordar que los actuales docentes fueron alumnos en un sistema educativo en el que la función de la evaluación se basaba en la medición y la acreditación, lo que supone años de experiencias evaluativas ancladas en una evaluación tradicional (Clark y Rust, 2006; López-Pastor y Pérez-Pueyo, 2017).

Esto último nos lleva a reflexionar sobre aquello que piensan los propios alumnos sobre la evaluación. Acercamientos actuales a la investigación sobre las concepciones que tienen los estudiantes, señalan que los alumnos piensan que la evaluación debería servir para aprender, en tanto que un buen feedback permite reflexionar sobre su progreso y saber cómo mejorar (Brown, 2011; Canabal y Margalef, 2017; Hidalgo y Murillo, 2017). Adicionalmente, una concepción surgida de prácticas de evaluación formativas ayuda a que los alumnos se impliquen, mejoren su autorregulación y posibilita mejores resultados (Monforte y Farías, 2013; Segers y Tilema, 2011). Sin embargo, otras investigaciones muestran que la realidad es que los alumnos descubren que se les evalúa para medir sus aprendizajes y nada más (Rueda y Torquemada, 2008).

A nivel de estrategias, analizamos aquellas decisiones que definen sus prácticas evaluativas, caracterizándolas respecto a quién debe evaluar, cuándo debe hacerlo y qué instrumentos y actividades va a realizar para valorar los resultados de aprendizaje. A diferencia de las concepciones, en esta dimensión encontramos que la tendencia mayoritaria se dirige hacia una evaluación auténtica. Además, también analizamos para esta dimensión, cómo describían los docentes sus pruebas de evaluación actuales. Los resultados nos muestran que gran parte de las características de la evaluación auténtica trabajadas en la Guía GAPPISA, eran utilizadas por los docentes en el diseño de nuevas pruebas. Así pues, creemos que la influencia de la formación ha tenido sus resultados positivos, especialmente en las prácticas evaluativas como señalamos en artículos anteriores (Álvarez-Bernárdez y Monereo, 2016; Mejías y Monereo, 2018; Monereo, Castelló et al., 2009; Monereo y Morcillo, 2011).

Pero también, como en el caso anterior, encontramos tensiones internas, por ejemplo, entre el instrumento más utilizado y el principal agente de evaluación. Los docentes utilizan el portafolio para recoger las evidencias de los aprendizajes de los alumnos y poder evaluarlas, pero no permiten que los alumnos también las evalúen y reflexionen sobre su aprendizaje. Por tanto, si no promueven que los alumnos sean agentes activos en su evaluación, el portafolio deja de ser un instrumento que posibilite la autorregulación (Barberà, Gewerc y Rodríguez, 2016). Aunque pensamos que esta incorrecta interpretación del instrumento portafolio puede estar derivada de una falta de conocimiento sobre este, el hecho de que los docentes declare que el principal objeto de la evaluación son los resultados y no cómo los alumnos aprenden, reafirma la idea de que los docentes todavía sostienen representaciones hacia una evaluación tradicional. Sin embargo, el hecho de que la mayor parte de los profesores en sus pruebas actuales introdujeran aspectos como compartir los criterios de evaluación y proponer tareas de autoevaluación y coevaluación, da cuenta del interés por orientar la evaluación al aprendizaje.

En cuanto a los sentimientos, en nuestro estudio, valoramos las emociones dentro del posicionamiento docente sobre evaluación como aquellas relacionadas con las prácticas y creencias sobre evaluación. No analizamos por tanto si emiten valoraciones según tiendan a un afecto positivo o negativo, sino que nos centramos

en conocer cuál es su satisfacción general y cuáles son sus principales preocupaciones, aspecto en el que incluimos los incidentes que declaran tener. Además, valoramos la importancia de trabajar a través del reconocimiento de aquello que sucede no aquello que debería suceder, para que los docentes no se sientan juzgados. Esta decisión se justifica porque sabemos que las emociones juegan un papel importante en la puesta en acción del rol profesional (Van Veen y Lasky, 2005) y en la reconstrucción de sus representaciones identitarias (Darby, 2008).

El sentimiento imperante en los resultados es el de satisfacción, motivación y ganas de compartir. Vemos este resultado positivo y promotor de cambio, puesto que, por un lado, la motivación y, especialmente, el hecho de compartir con la comunidad educativa facilita la sostenibilidad de los procesos de innovación (Liesa et al., 2018). Y, por otro lado, porque promueve mejoras en la relación que tienen los alumnos con la evaluación, en tanto que se permite su implicación y se crean escenarios seguros. Así es, estudios actuales sugieren que las emociones que expresan los docentes tienen tanta importancia para sus alumnos como las prácticas educativas que proponen (Becker, Goetz, Morger y Ranellucci, 2014).

Por último, hacemos referencia a los incidentes críticos que explicaban los docentes. Lo primero que debemos recordar es que algunos explicaron sucesos sobre evaluación que impactaron positivamente en su proceso de innovación. Estas experiencias son un aliciente para continuar con el cambio de prácticas evaluativas, y el hecho de que se aún sean capaces de recordarlos y los narren por encima de los negativos, puede verificar esta hipótesis. A la vez, esta declaración viene apoyada por el hecho de que muchos están bastante satisfechos respecto a la evaluación y, por supuesto, con el hecho de que hayan obtenido resultados positivos durante la formación sobre evaluación.

Respecto a los incidentes negativos, el hecho de que hagan referencia a situaciones sobre cómo aplicar la evaluación y también sobre lo que significa evaluar y lo que esta innovación provoca, nos muestra que la introducción de una innovación centrada en evaluación auténtica incide más allá de los cambios en las prácticas que los docentes puedan realizar de forma individual. Cambiar la

evaluación tiene implicaciones más allá, como en las metodologías y la cultura de centro (Moreno, 2012). De las preocupaciones declaradas, emergen temas como la coherencia entre la evaluación y la práctica, la cultura de centro sobre evaluación o la poca preparación de los alumnos, lo cuál nos lleva a concluir que cambiar la evaluación tiene implicaciones en los procesos de enseñanza y aprendizaje. Y, especialmente, que desplazarse hacia una evaluación auténtica genera movimientos internos dentro de las creencias y prácticas del rol docente, así como también en el contexto en que este participa.

Estos incidentes podríamos interpretarlos en relación con uno de los pocos estudios realizados sobre incidentes críticos específicos en evaluación, donde Monereo (2019) como resultado de un análisis de un centenar de respuestas de docentes describió cinco bloques principales en los que situarlos: incidentes sobre la dificultad de compartir criterios de evaluación con los compañeros, incidentes relacionados con conflictos éticos, incidentes que se relacionan con la calificación interpretada como fracaso y los incidentes sobre las dificultades de evaluar competencias y expectativas desajustadas. Lo destacable de este análisis es el apartado de estrategia de afrontamiento que acompaña cada bloque.

Pasamos a resumir, a la luz de todas estas reflexiones, cómo describimos el posicionamiento de los docentes sobre evaluación después de participar en la formación GAPPISA. Respecto a qué piensan, hacen y sienten, la mayoría de los docentes declaran que tienen como objetivo al evaluar ofrecer ayudas para que los alumnos se corrijan y que utilizan el portafolio como instrumento principal. Sin embargo, los docentes no tienen como objeto de evaluación las estrategias de aprendizaje de los alumnos, aspecto consustancial al hecho de querer ofrecer ayudas para que los alumnos mejoren. Tampoco declaran que el principal agente de la evaluación sean los propios alumnos o, al menos, mayoritariamente alumnos y docentes. Estas dificultades se reafirman cuando los profesores confiesan que su mayor preocupación es saber aportar evidencias útiles para que los propios alumnos puedan mejorar. Sabemos que las representaciones que conforman la identidad son difíciles de reescribir y que una construcción es más fuerte cuanto más ha sido sostenida en el tiempo.

Atendiendo ahora al segundo objetivo del estudio consistente en explorar la relación entre los posicionamientos sobre evaluación y la calidad y permanencia de las pruebas diseñadas durante la formación, los resultados parecen indicar que ni la calidad de la prueba ni el tiempo desde que se aplicó parecen influir en la permanencia ni en la satisfacción que el docente siente sobre su manera de evaluar. Si hablamos de permanencia, en primer lugar, el hecho de que la mayoría de los docentes declaren seguir utilizando su prueba y, especialmente, diseñar más pruebas siguiendo los presupuestos de la Guía GAPPISA nos muestra que, a pesar de las dificultades en la aplicación o los incidentes surgidos después de esta, las pruebas competenciales, formativas e inclusivas son útiles y necesarias. En segundo lugar, aunque las pruebas que diseñan mayoritariamente no cumplan todos los criterios, pensamos que no quiere decir que no apliquen pruebas auténticas y competenciales, sino que utilizan los criterios en diferentes actividades de evaluación y, seguramente, de enseñanza.

Las implicaciones a nivel práctico de estos resultados también nos parecen merecedoras de una reflexión. El hecho de que el posicionamiento sobre evaluación tienda hacia una evaluación auténtica pero todavía no esté fuertemente consolidado en ella, pensamos que es positivo por dos motivos: uno, nos demuestra que los cambios son progresivos y situados en la realidad personal y profesional de cada profesor, y dos, se pueden identificar qué aspectos todavía son confusos y trabajar sobre ello.

Uno, cambiar la posición evaluadora no es rápido ni sencillo, se requiere una reflexión sobre la práctica y sobre las motivaciones constante, en definitiva, ser competente en evaluación (Sanahuja y Sánchez-Tarazaga, 2018) y hacerlo desde una visión consensuada y situada a nivel de centro, con la ayuda de la comunidad de práctica (Liesa et al., 2018). Por ello, aunque la mayoría de los docentes tiendan hacia una visión auténtica después de la formación y a la introducción de nuevas prácticas, de forma sostenida en al menos dos años, se requiere un cambio mayor a nivel individual y también a nivel contextual.

Y dos, identificar qué concepciones, estrategias y/o sentimientos generan contradicción es necesario para reconstruir las posiciones evaluativas. Quizás

cambiar el método de evaluación no es suficiente para que la práctica sea sostenible (Monereo, Badia et al., 2009), pero sí un primer paso para iniciar el proceso de transformación a través de cambios en productos palpables (Mejías y Monereo, 2018) y en el discurso sobre evaluación (Badia y Chumpitaz-Campos, 2018) y continuar reflexionando, como estrategia formativa sobre qué aspectos debe trabajar para construir un posicionamiento estratégico, consistente y estable. Por este motivo defendemos, a partir de estos resultados, el diseño de asesoramientos centrados en la identidad, docente iniciados desde la práctica real en el aula, y centrados en cómo construir un posicionamiento sólido sobre evaluación auténtica como estrategia para hacer los cambios más sostenibles.

Y, como puerta de entrada a este análisis de la propia identidad, creemos que trabajar en torno a los incidentes críticos es una buena propuesta. Las formaciones a partir de los incidentes críticos se han demostrado realmente eficaces en su cometido de hacer emerger en los docentes nuevas posiciones que les permitan afrontar diversas situaciones emocionalmente críticas (Bilbao y Monereo, 2011; Monereo, Weise et al., 2013; Nail et al., 2012; Weise y Sánchez-Busqués, 2013). Sabemos que estos incidentes obligan al docente a replantearse ciertas representaciones sobre su identidad para analizar el problema desde otras perspectivas que puedan darle respuesta. Pensamos que trabajar sobre ello permitirá a los docentes continuar creciendo en su proceso de cambio hacia una evaluación auténtica y superar las disonancias internas restantes para la sostenibilidad de sus prácticas de evaluación.

3.5. Conclusiones, Limitaciones y Prospectiva

En conclusión, los resultados nos muestran que los docentes, después de realizar la formación sobre evaluación auténtica, Proyecto GAPPISA y, a pesar de haber pasado de uno a tres años después de su participación, declaran tener un posicionamiento sobre evaluación con una tendencia clara hacia las características que definen una evaluación auténtica. Así pues, realizar un cambio en las prácticas de evaluación provoca replantearse el propósito de esta, así como introducir estrategias que pueden mejorar los procesos de enseñanza y aprendizaje.

Así pues, como ha ocurrido en otras experiencias formativas (López-Pastor y Pérez-Pueyo, 2017; Sondergeld et al., 2010; William et al., 2004), nuestra propuesta resultó ser un buen medio para promover cambios en las prácticas evaluativas especialmente, pero, también, de forma iniciática, creemos que también en las concepciones sobre evaluación y sus sentimientos asociados, tal y como demuestra el hecho, por ejemplo, de que diversos incidentes hicieran referencia a la puesta en práctica de la evaluación auténtica o a que los docentes estén preocupados por el tipo de evidencias que recogen para que su evaluación sea realmente formativa.

Sin embargo, encontramos limitaciones en nuestro estudio que precisan considerarse para nuevas investigaciones. Por un lado, la manera dicotómica de describir los perfiles y, por otro lado, el estudio exclusivamente declarativo y poco contextualizado de los posicionamientos.

Por un lado, pensamos que es necesario un análisis no de las tendencias sino de los perfiles, en el caso de tener una muestra más representativa con el fin de encontrar posibles patrones. Hemos descrito las tendencias generales en el posicionamiento docente sobre evaluación que los docentes declaran a partir de contraponer dos tipos de evaluación, con el objetivo de que los docentes puedan clarificar su discurso. Sin embargo, sabemos que las concepciones sobre lo que significa evaluar son complejas, situadas y no excluyentes (Coll y Remesal, 2009). Se refleja este hecho en el análisis de los descriptores de cada una de las dimensiones del posicionamiento que hemos analizado, en el que siempre se daban tensiones entre ellos, aspecto común a todos los casos. Así pues, es preciso describir el perfil de cada maestro en profundidad y, para ello, realizar entrevistas individuales y grupales, analizar documentos, etc. En definitiva, recoger más datos que nos permitan profundizar en cada caso para describirlo de forma más compleja y situada. Además, en este sentido, creemos necesario revisar el cuestionario de posicionamiento docente sobre evaluación profundizando en las diversas variables, especialmente en la dimensión de emociones, para la cual podríamos servirnos del estudio de Badia, Meneses y Monereo (2014).

Por otro lado, sin duda resulta positivo que los docentes sepan declarar lo que piensan, lo que valoran que hacen y lo que creen que sienten, pero, ni decirlo quiere decir hacerlo, ni hacerlo quiere decir que se convierta en norma para todas las situaciones, como ya hemos apuntado cuando hablábamos de la brecha existente entre concepciones y prácticas en evaluación (Hidalgo y Murillo, 2017). La diferencia entre la identidad declarada y la identidad en acción, siendo ésta última múltiple y discontinua (Akkerman y Meijer, 2011; Monereo y Monte, 2011), nos muestra nuevas posibilidades de investigación. El estudio que hemos realizado se basa en la primera, en aquello que confiesan los profesores sobre sus creencias, acciones y sentimientos y partir de la voz de los protagonistas es necesario y, seguramente, supone el primer paso, pero parece no ser suficiente.

Consecuentemente, si de verdad queremos trabajar sobre su identidad docente para transformarla y que las prácticas sean sostenibles en el tiempo, debemos analizar también cómo es esta puesta en práctica. Es necesario pues examinar a qué finalidades realmente responden las acciones en el aula, con qué componentes se describen sus actividades evaluativas diarias y qué emociones y reacciones emergen de esta puesta en práctica. Obviamente la metodología y, especialmente, la muestra en la distancia, temporal y espacial, que definieron en su día la investigación, no posibilitó esta intención. Pero señalamos esta propuesta para futuras investigaciones que, avaladas por la literatura más reciente respecto a la identidad de los docentes, podrían realizarse a la luz de la teoría dialógica del yo (Hermans, 2001, 2009, 2014), centrada en cuál es el diálogo de las voces sobre evaluación que activa el docente para resolver sus incidentes y generar nuevas posiciones.

CAPÍTULO 4. Conclusiones

En este capítulo vamos a describir y justificar de qué manera los estudios realizados se interrelacionan para dar respuesta a los objetivos planteados en la investigación, qué aportación realizan a la literatura existente sobre la evaluación auténtica y sobre la identidad docente y la posición evaluadora y cuáles son las implicaciones que tienen especialmente para la mejora de las prácticas de evaluación mediante la formación docente. Y, para acabar, resumiremos las limitaciones encontradas y propondremos nuevas líneas de investigación.

El propósito de la investigación ha sido en todo momento mejorar la evaluación de profesores y maestros a través de una formación que impactara en su posicionamiento docente sobre evaluación como medida de sostenibilidad en los cambios evaluativos. Este objetivo nació en un contexto español en el que los docentes reclamaban estrategias para poder adaptar la evaluación a las nuevas exigencias del currículum, que introducía con mayor fuerza por entonces, el enfoque de enseñanza por competencias. La evaluación orientada al aprendizaje empezó a predominar en este contexto, proponiéndose como solución al gran reto que la enseñanza competencial proponía: la autorregulación de aprendizajes complejos por parte de los alumnos. Paralelamente, empezó a emerger la evaluación auténtica que, además de incluir notables características de la evaluación para el aprendizaje, proponía una evaluación competencial en tanto que enfocaba sus prácticas a contextos situados y realistas, mediante la resolución de problemas complejos, acciones colaborativas y regulación mediada durante la resolución. De esta manera, nació el objetivo de formar a docentes para que transformaran su evaluación tradicional en una evaluación auténtica.

En este contexto, de forma transversal a estas nuevas necesidades formativas sobre evaluación, estaba presente una preocupación que psicólogos educativos y pedagogos han investigado durante mucho tiempo, esto es, conseguir que los resultados de las formaciones docentes se transfirieran a la práctica real de los centros educativos con eficacia y permanencia. Este desafío, de alta complejidad debido a la naturaleza del objeto de estudio condicionado por tantas variables, empezó a investigarse desde una nueva perspectiva, la de la identidad docente.

Coincidiendo con el aumento de literatura sobre identidad docente desde el ámbito de la psicología de la educación, nacieron nuevos estudios centrados en examinar qué relación había entre ésta y las actuaciones de los profesores, en tanto que se descubrió que, a la práctica, los profesionales educativos activaban diferentes posiciones según el diálogo interno que sostenían y el contexto en el que se desarrollaban. Es decir, estudios sobre cómo la comprensión de la identidad docente podría ayudarnos a mejorar el asesoramiento educativo. De esta manera, nació la posibilidad de comprobar si a través del conocimiento de estas posiciones, en nuestro caso, evaluativas, se podía mejorar la formación y conseguir que el cambio en evaluación fuera más sostenible en el tiempo, asegurado por un docente competente en evaluación que es capaz de afrontar los retos e incidentes que le pudieran surgir.

Este recorrido por el contexto del que nace la investigación es un resumen para explicar de forma más directa la naturaleza del objetivo de la investigación, más desde un análisis fenomenológico y no teórico, dado que la literatura al respecto la hemos desarrollado en los capítulos anteriores. Sin embargo, nos parece interesante y necesario hacerlo para situar en contexto los temas que se han desarrollado en esta investigación y, especialmente, poder entender las argumentaciones que se realizarán a continuación.

Para responder al propósito de la investigación, se plantearon dos estudios, con los siguientes objetivos: un primer estudio centrado en analizar los resultados y la eficacia de una formación sobre evaluación auténtica realizada a docentes que quería mejorar su evaluación; y un segundo estudio dirigido a identificar y describir las posiciones evaluativas de los docentes que habían participado en la formación. De esta manera, por un lado, demostraríamos si la formación sobre evaluación auténtica, a nivel de contenido y de forma, obtenía buenos resultados para poder replicarla, respondiendo así a la necesidad detectada en el contexto educativo de alfabetizar a los docentes en evaluación auténtica. Y, por otro lado, analizar si los docentes que habían participado en la formación sostenían posiciones evaluativas cercanas a un enfoque auténtico, lo que nos permitiría declarar que la formación tuvo un impacto en la identidad docente. También, identificando y describiendo las

posiciones evaluativas podríamos averiguar posibles patrones existentes y/o dificultades comunes que pudieran abrir paso a nuevas formaciones.

Respecto al primer objetivo, podemos confirmar que se ha cumplido: la calidad de las pruebas de evaluación respecto a su autenticidad, diseñadas por los docentes gracias a la formación, da cuenta de su efectividad, tanto en lo que se refiere a su programación pedagógica como a los contenidos desarrollados sobre evaluación auténtica. El hecho de que los profesores diseñaran pruebas más competenciales, que incluyen aspectos formativos y que son más inclusivas, significa también que el marco que hemos desarrollado sobre evaluación auténtica está en sintonía con los objetivos de enseñanza de los docentes y les aporta las herramientas necesarias para mejorar su evaluación dentro del panorama educativo actual.

Respecto al segundo objetivo, también podemos verificar su cumplimiento, en tanto que de forma mayoritaria los docentes declaran tener un posicionamiento sobre evaluación que tiende a un enfoque de evaluación auténtica. Sin embargo, no hemos hallado patrones comunes que nos permitan describir e identificar posiciones prototípicas sobre evaluación; por el contrario, cada docente nos mostraba su singularidad en la manera de posicionarse frente la evaluación. Pensamos que este resultado, como ya indicamos en su momento, puede reinterpretarse desde la literatura sobre identidad docente desarrollada en la que existen identidades-del-yo situadas, y, por lo tanto, una multiplicidad de posiciones, en este caso sobre evaluación, tantas como realidades educativas.

Con relación a los resultados del primer estudio, podemos decir que, para la literatura sobre evaluación auténtica, aportan nuevos datos sobre el diseño de prácticas de evaluación auténtica, específicamente, a la literatura sobre transformación de exámenes en pruebas de evaluación auténticas (Álvarez-Bernárdez y Monereo, 2016; Monereo, Castelló et al., 2009; Monereo y Morcillo, 2011; Villarroel, Boud, Bloxham, D. Bruna y C. Bruna, 2019), un tema en el que no hay tanta literatura existente como sí la hay en relación a otros instrumentos identificados de evaluación auténtica, como son los portafolios o las rúbricas de evaluación. Nos parece importante debido a la realidad docente en la que nos

encontramos, donde el examen tiene un puesto predominante en los instrumentos de evaluación más utilizados en todas las etapas de educación obligatoria.

Estos resultados también aportan evidencias para la práctica en el ámbito de la formación docente. Primero, destacando la necesidad de partir de la zona de desarrollo de los profesores, es decir, de sus prácticas más habituales, para conseguir transformaciones reales. Segundo, corroborando que la figura del formador o del asesor educativo debe responder al rol, no de transmisor de unos saberes, sino al de acompañar durante el aprendizaje y la puesta en práctica, para poder evaluar conjuntamente los éxitos conseguidos; aspecto que conecta con nuevas estrategias formativas efectivas como el coaching docente o la docencia compartida (Johnson et al., 2017; Kraft et al., 2018; Oller et al., 2018). Tercero, la promoción de una comunidad de práctica a través de estrategias como la creación de una base de pruebas de evaluación a la que pueden acceder todos los participantes o la creación de comisiones sobre evaluación coordinadas por un docente responsable, guiado por los asesores; aspectos que también se justifican con investigaciones que apuntan que, para la sostenibilidad de los cambios, es necesario transformar el centro en "comunidades de aprendizaje profesionales" (Liesa et al., 2018).

Con relación a los resultados del segundo estudio, de carácter exploratorio, podemos confirmar que inicia una nueva línea de investigación centrada en analizar las posiciones evaluativas de los profesores, una temática emergente y sin apenas investigación al respecto, al menos desde la perspectiva de la identidad docente. Una línea de investigación que se ha apoyado en los estudios anteriores en los que se comprobaba si el cambio de prácticas en evaluación había incidido de alguna manera en las concepciones docentes sobre evaluación (Maclellan, 2004; Rueda y Torquemada, 2008; Rodríguez, 2010), con la pretensión de ir más allá y estudiar las posiciones identitarias. Alguna investigación ha empezado a acercarse a este enfoque (Badia y Chumpitaz-Campos, 2018) y contamos con otras investigaciones que lo apoyan, especialmente, en su estudio a través de incidentes críticos (Nail et al., 2012; Monereo, Badia et al., 2009; Monereo, Weise et al., 2013). Por último, pensamos que tiene implicaciones en la práctica, en tanto que el estudio en profundidad de las posiciones evaluativas y los incidentes críticos en evaluación

aporta estrategias formativas para la formación docente, como desarrollaremos más adelante.

Sin embargo, aunque las aportaciones que se deducen de los resultados son positivas y variadas, los estudios tienen limitaciones que es preciso remarcar, especialmente para mejorar en futuras investigaciones. Estas son: uno, limitaciones relativas al asesoramiento a nivel de centro; dos, limitaciones relativas a la descripción del posicionamiento docente sobre evaluación; y, tres, limitaciones relativas al análisis de los cambios en el posicionamiento docente sobre evaluación.

En primer lugar, si queremos conseguir un impacto en las prácticas evaluativas de los docentes de forma que estas sean sostenibles en el tiempo es necesario, no únicamente cambiar las prácticas, sino también la cultura evaluativa de centro. Como hemos apuntado anteriormente en la literatura analizada, es necesario que los docentes reflexionen conjuntamente sobre sus prácticas y tomen decisiones consensuadas a nivel de centro, en las que todo el equipo educativo esté implicado y colabora en la construcción de una cultura compartida (Martínez, 2013; Sondergeld et al., 2010; Xarxa de Competències Bàsiques, 2019). Nos acercamos a este objetivo a través de la figura del coordinador, encargado de promover y dirigir una comisión sobre evaluación y, también, con el seminario de formación inicial y la comunicación de los resultados de la formación, momentos en los que iniciamos procesos de reflexión sobre la evaluación a nivel de centro. Sin embargo, valoramos que, a efectos de conseguir la permanencia de la evaluación auténtica en el tiempo, la formación podría haberse extendido más allá del contexto de aula, proponiendo actividades que aseguraran un proceso colaborativo de transformación de la cultura evaluativa a nivel de centro.

En este proceso de apertura de la formación, además, ha sido una limitación no haber considerado los datos de los resultados de la implementación de las pruebas de evaluación auténtica. El análisis de los cambios que se hayan podido producir en cómo el docente diseña su enseñanza y en cómo los alumnos enfocan su aprendizaje, hubiera facilitado una comprensión más profunda de los efectos de la evaluación auténtica y, por tanto, de los resultados de la formación. Sin embargo, debido al carácter no presencial de la formación, el acceso a la muestra resultó

difícil. No obstante, debemos tener presente estas variables para futuras investigaciones.

En segundo lugar, una de las limitaciones en el momento de analizar el posicionamiento docente sobre evaluación consistió en encontrar literatura que nos permitiera establecer un marco de referencia. Como hemos explicado anteriormente, la literatura sobre la identidad docente y el estudio de las posiciones-del-yo, es todavía joven. Así pues, en la descripción de las dimensiones de la identidad, para las concepciones nos hemos podido basar en estudios descriptivos sobre las representaciones de los docentes en su rol de evaluadores y, para la dimensión de estrategias, además también del estudio de las concepciones sobre evaluación de los docentes, especialmente, en la literatura sobre cómo diseñar e implementar la evaluación. Sin embargo, para la dimensión de emociones y sentimientos que los docentes sostienen respecto a la evaluación hemos partido, mayoritariamente, de estudios en los que se trataba esta temática de forma transversal. Esto ha representado cierta limitación en el momento de interpretar esta dimensión y es necesario profundizar en este sentido, apoyados por trabajos que de esta dimensión se ha hecho desde otras temáticas (Badia, 2014; Badia et al., 2014).

En tercer lugar, para poder describir en profundidad las posiciones evaluativas, pensamos que se deben tener en cuenta más variables de las consideradas. Esta ampliación podríamos conseguirla, además de con una investigación más exhaustiva de la literatura, a través del estudio cualitativo de dicho posicionamiento desde una perspectiva dialógica, indagando sobre cuáles son las voces que los docentes utilizan para describir su evaluación, especialmente de manera previa a la formación. Y, precisamente, esta ha sido la tercera limitación, relativa al análisis de los cambios en el posicionamiento docente sobre evaluación: al no disponer de un análisis previo en profundidad de cómo los docentes construían su posición evaluativa antes de la formación, no podemos contrastar e identificar qué características son las que ha variado gracias a la influencia de la formación.

La identificación de estas limitaciones, en su conjunto y por separado, nos ayudan a determinar de qué manera pueden enfocarse nuevas líneas de investigación, en

los dos ámbitos estudiados, la formación docente sobre evaluación y el estudio de las posiciones evaluativas de los docentes. En este caso, creemos necesario seguir interrelacionando los dos ámbitos para poder responder con mayor profundidad al propósito planteado: mejorar la evaluación de profesores y maestros a través de una formación que impacte en su posicionamiento docente sobre evaluación como medida de sostenibilidad en los cambios evaluativos.

Una de las futuras investigaciones que destacamos por su especial interés para el ámbito de la evaluación, es el estudio de las posiciones evaluativas de los docentes desde la acción, es decir, examinar las posiciones-del-yo a la luz de la teoría dialógica de Hermans (2001, 2009, 2014) y los primeros acercamientos que se han hecho en este sentido (Badia y Chumpitaz-Campos, 2018). Así pues, comprobado que el estudio de la posición declarada es un primer paso, es necesario examinar cuál es el diálogo de las voces que el docente activa sobre evaluación para enfrentarse a los retos e incidentes surgidos de la innovación en evaluación. Este estudio implicaría, a nivel metodológico, el uso de instrumentos cualitativos de recogida y análisis de datos, como entrevistas, observación no participante durante las prácticas evaluativas y/o análisis de los instrumentos de evaluación que utilicen, y, cómo no, la descripción de los incidentes críticos sobre evaluación. También, debería contemplar momentos pre y post formación para poder comparar los cambios acaecidos. De esta manera, podríamos conocer el nivel de competencia en evaluación del docente, estudiando cuáles son sus fortalezas y debilidades en el momento de posicionarse ante la evaluación y cómo la formación ha incidido en estas.

Obviamente, de este estudio se desprende otra línea consistente en actualizar la formación a la luz de estos nuevos conocimientos sobre las posiciones evaluativas y los incidentes en evaluación. Esta nueva propuesta partiría de la concepción de que cada docente se posiciona ante la evaluación de una manera singular y contextualizada y, por lo tanto, requiere de un acompañamiento mucho más individualizado y situado en su contexto de centro. Así pues, estrategias como el "coaching" entre docentes, la docencia compartida o el asesor como generador de contextos de cambio y conductor de las comunidades profesionales de práctica, que se han demostrado efectivas en la transferencia del aprendizaje de los

docentes al aula (Badia y Monereo, 2004; Johnson et al., 2017; Kraft et al., 2018; Oller et al., 2018) podrían constituir la base de la nueva propuesta. Por supuesto, el análisis de los incidentes críticos mediante talleres prácticos de exposición de estas situaciones reales a través de dramatizaciones o narraciones y su resolución colaborativa, serían también actividades nucleares de la formación (Del Maestro y Monereo, 2014; Fernández y Marrero, 2013; Monereo, Weise et al., 2013; Weise y Sánchez-Busqués, 2013).

En conclusión, con las investigaciones realizadas se inicia un nuevo campo de investigación sobre el rol de los profesores como evaluadores dentro de su identidad docente y se realizan recomendaciones claras sobre cómo deben ser las formaciones docentes para conseguir mejorar su competencia en evaluación desde el modelo de evaluación auténtica de competencias, la cuál permite evaluar de forma competencial, formativa e inclusiva los aprendizajes de los alumnos. Esperamos pues, poder seguir investigando en esta línea con el objetivo de diseñar propuestas de formación continua de calidad, basadas en los hallazgos científico, que psicólogos educativos y pedagogos puedan implementar en los centros con resultados exitosos en lo que se refiere a los aprendizajes de los alumnos.

REFERENCIAS

- Abdul, S. S., Venville, G. y Chapman, A. (2009) Classroom Assessment: Juxtaposing Teachers' Beliefs with Classroom Practices. *Australian Association For Research In Education: International Education Research Conference*. Recuperado de: <https://www.aare.edu.au/data/publications/2009/abd091051.pdf>
- Adell, J., Mengual-Andrés, S. y Roig-Vila, R. (2015). Webquest: 20 años utilizando Internet como recurso para el aula. *EDUTECA, Revista Electrónica de Tecnología Educativa*, 52.
- Ahumada, P. (2005). La evaluación auténtica: un sistema para la obtención de evidencias y vivencias de los aprendizajes. *Perspectiva Educativa, Instituto de educación PUCV*, 45(1), 11-24.
- Ainscow, M. (2012). Haciendo que las escuelas sean más inclusivas: lecciones a partir del análisis de la investigación internacional. *Revista de Educación Inclusiva*, 5(1), 39-49.
- Akkerman, S.F. y Meijer, P.C. (2011). A dialogical approach to conceptualizing teacher identity. *Teaching Education*, 27(2), 308-319.
- Alcañiz, V. y Cervera, D. (2014). Evaluaciones externas, mucho más que resultados. Una mirada centrada en PISA. *Revista Avances en supervisión educativa*, 21. Recuperado de <https://avances.adide.org/index.php/ase/article/view/84>
- Alderete, A. y Edith, K. (2018). Evaluación del Desempeño y Auténtica en el Modelo por Competencias en Secundaria: Un Estudio Mixto. *REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 16(3), 103-122.
- Alfageme, M. B., Miralles, P. y Monteagudo J. (2010). Diseño y validación de un instrumento sobre evaluación de la geografía y la historia en educación secundaria. *Enseñanza de las ciencias sociales*, 10, 51-63.

- Alsina, J. (Ed.) (2013). *Rúbricas para la evaluación de competencias*. Barcelona: OCTAEDRO.
- Álvarez-Bernárdez, P. y Monereo, C. (2016). Formación del profesorado de ELE y evaluación. Construcción de una guía para autenticar pruebas de español lengua extranjera. *Porta Linguarum*, 25(1), 163-177.
- Álvarez, I.M. (2005). Evaluación como situación de aprendizaje o Evaluación auténtica. *Perspectiva Educativa, Formación de Profesores*, 45, 45-68.
- Álvarez, I.M. (2009). Evaluar para contribuir a la autorregulación del aprendizaje. *Electronic Journal of Research in Educational Psychology*, 7(3), 1007-1030.
- Álvarez, J.M. (2001). *Evaluar para conocer, examinar para excluir*. Madrid: Morata.
- Ashford-Rowea, K., Herrington, J. y Brown, C. (2014). Establishing the critical elements that determine authentic assessment. *Assessment & Evaluation in Higher Education*, 39(2), 205-222.
- Avery, P.G., Freeman, C. y Carmichael, D. (2012) Developing Authentic Instruction in the Social Studies. *Journal of Research in Education*, 12(1), 50–56.
- Badia, A. (2014). Emociones y sentimientos del profesor en la enseñanza y la formación docente. En C. Monereo (Coord.). Enseñando a enseñar en la universidad (pp. 62-90). Barcelona: Octaedro/ICE-UB.
- Badia, A. y Chumpitaz-Campos. L. (2018). Teachers learn about student learning assessment through a teacher education process. *Studies in Educational Evaluation*, 58, 1-7.
- Badia, A., García, C. & Meneses, J. (2017). Approaches to teaching online: Exploring factors influencing teachers in a fully online university. *British Journal of Educational Technology*, 48(6), 1193-1207.

- Badia, A. y Monereo, C. (s.f.). La perspectiva dialógica en la construcción de la identidad humana a lo largo de la vida. Módulo 3 de la asignatura del grado de Psicología: "Aprendizaje a lo largo de la vida". Editorial EDIUOC (en prensa)
- Badia, A., Meneses, J., y Monereo, C. (2014). Affective dimension of university professors about their teaching: An exploration through the semantic differential technique. *Universitas Psychologica*, 13(1), 161-173.
doi:10.11144/Javeriana.UPSY13-1.adup
- Badia, A., y Monereo, C. (2004). La construcción de conocimiento profesional docente. Análisis de un curso de formación sobre la enseñanza estratégica. *Anuario de Psicología*, 35, 47-70.
- Barberà, E. (2005). La evaluación de competencias complejas: la práctica del portafolio. *Educere*, 31(9), 497-504.
- Barberà, E., Gewerc, A. y Rodríguez, J.L. (2016). Portafolios electrónicos y educación superior en España: Situación y tendencias. *RED. Revista de Educación a Distancia*, 50(7). Recuperado de <https://revistas.um.es/red/article/view/271151>
- Barca, A., Almeida, L., Porto, A., Peralbo, M. y Brenlla, J.C. (2012). Motivación escolar y rendimiento: impacto de metas académicas, de estrategias de aprendizaje y autoeficacia. *Anales de psicología*, 28(3), 848-859.
- Barnes, N., Fives, H. y Dacey, C. (2017). U.S. teachers' conceptions of the purposes of assessment. *Teaching and Teacher Education*, 65, 107-116.
- Becker, S., Goetz, T., Morger, V. y Ranellucci, J. (2014). The importance of teachers' emotions and instructional behavior for their students' emotions - An experience sampling analysis. *Teaching and Teacher Education*, 43, 15-26.
- Beijaard, D., Meijer, P. y Verloop, N. (2004). Reconsidering research on teachers' professional identity. *Teaching and Teacher Education*, 20, 107-128.

- Beijaard, D., Verloop, N. y Vermjunt, J. D. (2000). Teachers' perceptions of professional identity: An exploratory study from a personal knowledge perspective. *Teaching and Teacher Education: An International Journal of Research and Studies*, 16(7), 749-764.
- Bilbao, G. y Monereo, C. (2011). Identificación de incidentes críticos en maestros en ejercicio: propuestas para la formación permanente. *Revista Electrónica de Investigación Educativa*, 13(1), 135-151.
- Black, P. (2015). Formative assessment – an optimistic but incomplete vision. *Assessment in Education: Principles, Policy & Practice*, 22(1), 161-177.
- Black, P. y Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (EAEA)*, 21, 5-31.
- Blanch, S., Bosco, A., Gimeno, X., González, N., Fuentes, M., Jariot, M., ... Trafí, L. (2011). Carpetas de aprendizaje en la educación superior: una oportunidad para repensar la docencia. Bellaterra: Servei de Publicacions, Universitat Atuònoma de Barcelona.
- Blanco, M.E. y González, I. (2010). Políticas evaluativas del profesorado de primaria, condicionantes de su actitud docente. *Bordón*, 62(1), 29-47.
- Blasco, J.E., Mengual-Andrés, S. y Roig-Vila, R. (2011). Propuestas de integración de las tecnologías de la información y comunicación para la mejora de las prácticas docentes. *International Journal of Developmental and Educational Psychology: INFAD. Revista de Psicología*, 3(1), 657-666.
- Boud, D. (2000). Sustainable Assessment: Rethinking Assessment for the Learning Society. *Studies in Continuing Education*, 22(2), 151–167.
- Brown, G. T. L. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. *Assessment in Education: Principles, Policy & Practice*, 11(3), 301–318.
<http://dx.doi.org/10.1080/0969594042000304609>

- Brown, G. T. L. (2006) Teachers' conceptions of assessment: Validation of an abridged instrument. *Psychological Reports*, 99, 166–170.
- Brown, G. T. L. (2011). Self-regulation of assessment beliefs and attitudes: A review of the students' conceptions of assessment inventory. *Educational Psychology*, 31(6), 731–748.
- Brown, G. T. L. y Michaelides, M. P. (2011). Ecological rationality in teachers' conceptions of assessment across samples from Cyprus and New Zealand. *European Journal of Psychology of Education*, 26(3), 319–337.
- Brown, G. T. L. y Remesal, A. (2012). Prospective teachers' conceptions of assessment: a cross-cultural comparison. *The Spanish journal of psychology*, 15(1), 75–89.
- Brown, G. T. L., Hui, S., Yu, F. y Kennedy, K. (2011). Teachers' conceptions of assessment in Chinese contexts: A tripartite model of accountability, improvement, and irrelevance. *International Journal of Educational Research*, 50(5-6), 307–320.
- Brown, G. T. L., y Harris, L.R. (2014). The future of self-assessment in classroom practice: Reframing self-assessment as a core competency. *Frontline Learning Research*, 3, 22-30. doi:10.14786/flr.v2i1.24
- Brown, S. (2015). Authentic Assessment: Using assessment to help students learn. *Relieve*, 21(2), 1-10. doi: <http://dx.doi.org/10.7203/relieve.21.2.7674>
- Brown, S. (2015). La evaluación auténtica: el uso de la evaluación para ayudar a los estudiantes a aprender. *RELIEVE*, 21(2), doi: <http://dx.doi.org/10.7203/relieve.21.2.7674>
- Canabal, C. y Margalef, L. (2017) La retroalimentación: la clave para una evaluación orientada al aprendizaje. *Profesorado. Revista de Currículum y Formación de Profesorado*, 21(2), 149-170.
- Cano, E. y Ion, G. (Eds.) (2016). *Innovative Practices for Higher Education Assessment and Measurement*. Hershey (Pennsylvania): IGI Global.

- Carretero, R., Liesa, E., Mayoral, P., y Mollà, N. (2008). El papel de la motivación de los asesores y profesores en el proceso de asesoramiento. *Profesorado. Revista de currículum y formación del profesorado*, 12,1-15.
- Clark, C.M. y Rust, E.O'C. (2006). Learning-Centered Assessment in a Teacher Education. *Studies in Educational Evaluation*, 32, 73-82.
- Coll, C. y Martín, E. (1996). La evaluación de los aprendizajes: una perspectiva de conjunto. *Signos. Teoría y práctica de la educación*, 18, 64-77. Recuperado de http://www.quadernsdigitals.net/datos_web/hemeroteca/r_3/nr_46/a_672/672.html
- Coll, C. y Onrubia, J. (2002). Evaluar en una escuela para todos. *Cuadernos de Pedagogía*, 318, 50-54.
- Coll, C. y Remesal, A. (2009) Concepciones del profesorado de matemáticas acerca de las funciones de la evaluación del aprendizaje en la educación obligatoria. *Infancia y Aprendizaje*, 32(3), 391-404.
- Coll, C., Barberà, E. y Onrubia, J. (2000). La atención a la diversidad en las prácticas de evaluación. *Infancia y Aprendizaje*, 23(90), 111-132. doi: [10.1174/021037000760087991](https://doi.org/10.1174/021037000760087991)
- Coll, C., Mauri, T. y Rochera, MJ. (2012). La práctica de evaluación como contexto para aprender a ser un aprendiz competente. *Revista Pofesorado*, 16(1), 49-59.
- Comas, N. (2019). Serveix la formació permanent del professorat per millorar els resultats educatius de l'alumnat?. Fundació Jaume Bofill, Ivàlua. Recuperado de: http://www.ivalua.cat/documents/1/12_06_2019_13_10_27_Que_funciona_14_Catala.pdf
- Comisión Europea (2007). Competencias clave para el aprendizaje permanente. Un marco de referencia europeo. *Luxemburgo, Comunidades Europeas*.

Recuperado de:

<https://www.mecd.gob.es/dctm/ministerio/educacion/mecu/movilidad-europa/competenciasclave.pdf?documentId=0901e72b80685fb1>

- Condemarín, M. y Medina, A. (2000). Evaluación auténtica de los aprendizajes: una propuesta para el cambio. En M. Condemarín y A. Medina (Ed.), *Evaluación de los aprendizajes. Un medio para mejorar las competencias lingüísticas y comunicativas* (pp. 13-32). Chile: MINEDUC
- Coombs, A., DeLuca, C., y LaPointe-McEwan, D. (2018). Changing approaches to classroom assessment: An empirical study across teacher career stages. *Teaching and Teacher Education, 71*, 134-144.
- Darby, A. (2008). Teachers' emotions in the reconstruction of professional self-understanding. *Teaching and Teacher Education, 24*, 1160-1172.
- Darling-Hammond, L. y Snyder, J. (2000). Authentic assessment of teaching in context, *Teaching and Teacher Education, 16*(5), 523–545.
- Dawson, P. (2017). Assessment rubrics: towards clearer and more replicable design, research and practice. *Assessment & Evaluation in Higher Education, 42*(3), 347-360.
- Day, C., Kington, A., Stobart, G. y Sammons, P. (2006). The personal and professional selves of teachers: stable and unstable identities. *British Educational Research Journal, 32*(4), 572-601.
- De la Orden, A. (2011). Reflexiones en torno a las competencias como objeto de evaluación en el ámbito educativo. *Revista Electrónica de Investigación Educativa, 13*(2), 1-21.
- Del Maestro, C. y Monereo, C. (2014). Incidentes críticos en los profesores universitarios de la PUCP. *RIES, Revista Iberoamericana de Educación Superior, 12*(5), 3-20.

- DeLuca, C. (2012). Preparing teachers for the age of accountability: toward a framework for assessment education. *Teacher Education Yearbook XXI: A Special Issue of Action in Teacher Education*, 34(5/6), 576–591.
- DeLuca, C., LaPointe-McEwan, D. y Luhanga, U. (2016). Teacher assessment literacy: a review of international standards and measures. *Educational Assessment, Evaluation and Accountability*, 28(3), 251-272.
- Deneen, C. y Brown, G. (2016). The impact of conceptions of assessment on assessment literacy in a teacher education program. *Cogent Education*, 3(1). doi: 10.1080/2331186X.2016.1225380
- Díaz, A. (2006). El enfoque de competencias en la educación. ¿Una alternativa o un disfraz de cambio?. *Perfiles Educativos*, 28(2), 7-36.
- Díaz, F. y Barroso, R. (2014). Diseño y validación de una propuesta de evaluación auténtica de competencias en un programa de formación de docentes de educación básica en México. *Perspectiva Educativa. Formación de Profesores*, 53(1), 36-56.
- Dolores, C. y García-García, J. (2017). Concepciones de profesores de matemáticas acerca de la evaluación vistas a la luz de la reforma educativa actual en México. *Revista Paradigma*, 28(1), 186-210.
- Doménech, F., y Fortea, M. (2015). Validez Factorial de un Cuestionario para Evaluar la Percepción Inicial de los Estudiantes Universitarios sobre la Evaluación de los Aprendizajes. *Electronic Journal of Research in Educational Psychology*, 13(3), 569-586.
- Duran, D. y Giné, C. (2017). La formación del profesorado para la educación inclusiva: Un proceso de desarrollo profesional y de mejora de los centros para atender la diversidad. *Revista Latinoamericana de Educación Inclusiva*, 5(2), 153-170.

- Duran, D. y Monereo, C. (2008). Una guia per l'anàlisi de proves d'avaluació des de la perspectiva PISA. *Revista de Didàctica de la llengua i la literatura*, 44, 101-116.
- Elórtogui, N., Medina, M. y Fernández, J. (2003) Los incidentes críticos en la formación y perfeccionamiento del profesorado de secundaria de ciencias de la naturaleza. *Revista interuniversitaria de formación del profesorado*, 17(1), 101-112
- Fernández, J. y Marrero, J. J. (2013). La estrategia del análisis teórico de la práctica de profesores de Secundaria en formación: los incidentes críticos. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 10, 694-708.
- Frey, B., Schmitt, V. y Allen, J. (2012). Defining Authentic Classroom Assessment. *Practical Assessment, Research & Evaluation*, 17(2), 1-18.
- Fuentes, T. y López, V.M. (2017). Evaluación auténtica, coevaluación y uso de las TIC en educación física: un estudio de caso de secundaria. *Infancia, Educación y Aprendizaje*, 3(2), 42-46.
- Galvis, R. V. (2007). De un perfil docente tradicional a un perfil docente basado en competencias. *Acción pedagógica*, 16(1), 48-57.
- García-Jiménez, E. (2015). La evaluación del aprendizaje: de la retroalimentación a la autorregulación. El papel de las tecnologías. *RELIEVE*, 21(2), doi: <http://dx.doi.org/10.7203/relieve.21.2.7546>
- Generalitat de Catalunya (2015). De l'escola inclusiva al sistema inclusiu. Una escola per a tothom, un projecte per a cadascú. Recuperado de: <http://ensenyament.gencat.cat/web/.content/home/departament/publicacions/colleccions/inclusio/escola-inclusiva.pdf>
- Gómez-Devís, M. B., Garcia-Raffi, J. V. y Saneleuterio, E. (2017). Language Education and Assessment Skills Development. Analysis of an Experience

in Teachers Training. *Revista de investigaciones y experiencias en Ciencias de la Educación*, 26, 132-145.

Gómez, C.J. y Miralles, P. (2013). Los contenidos de ciencias sociales y las capacidades cognitivas en los exámenes de tercer ciclo de educación primaria ¿Una evaluación en competencia?. *Revista Complutense de Educación*, 24(1), 91-121.

Gómez, C.J., Monteagudo, J. y Miralles, P. (2018). Conocimiento histórico y evaluación de competencias en los exámenes de Educación Secundaria. Un análisis comparativo España-Inglaterra. *Educatio Siglo XXI*, 36 (1), 85-106.

Goodman, G., Arbona, C. y Dominguez, R. (2008). High-Stakes, Minimum-Competency Exams: How Competent Are They for Evaluating Teacher Competence? *Journal of Teacher Education*, 59(1), 24–39.

Guasch, T., Guàrdia, L. y Barberà, E. (2009). Prácticas del portafolio electrónico en el ámbito universitario del Estado Español. *RED. Revista de Educación a Distancia*. Recuperado de:
<https://revistas.um.es/red/article/view/69611/67101>

Gulikers, J. T. M., Bastiaens, T. J. y Kirschner, P. A. (2004). A Five-Dimensional Framework for Authentic Assessment. *Educational Technology Research and Development*, 52(3), 67–86.

Gulikers, J. T. M., Bastiaens, T. J., & Kirschner, P. A. (2006). Authentic assessment, student and teacher perceptions: the practical value of the five dimensional-framework. *Journal of Vocational Education and Training*, 58, 337-357

Hamodi, C. y López, A.T. (2012). La evaluación formativa y compartida en la Formación Inicial del Profesorado desde la perspectiva del alumnado y de los egresados. *Psychology, Society, & Education*, 4(1), 99-112.

- Heritage, M. (2018). Assessment for learning as support for student self-regulation. *Australian Educational Researcher*, 45, 51-63.
- Hermans, H.J.M. (2001). The dialogical self: Toward a theory of personal and cultural positioning. *Culture & Psychology*, 7(3), 243–281.
- Hermans, H.J.M. (2009). The Construction of a Personal Position Repertoire: Method and Practice. *Culture & Psychology*, 7(3), 323–365.
- Hermans, H.J.M. (2014). Self as a Society of I-Positions: A Dialogical Approach to Counseling. *Journal of Humanistic COUNSELING*, 53, 134-159.
- Herrington, A. y Herrington, J. (2006). Authentic conditions for authentic assessment: aligning task and assessment. In A. Bunker & I. Vardi (Eds.), *Proceedings of the 2006 Annual International Conference of the Higher Education Research and Development Society of Australasia Inc (HERDSA): Critical Visions: Thinking, Learning and Researching in Higher Education*. Research and Development in Higher Education, Volume 29 (pp. 141-151). Milperra, NSW: HERDSA.
- Hidalgo, N. y Murillo, F.J. (2017). Las concepciones sobre el proceso de evaluación del aprendizaje de los estudiantes. *REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 15(1), 107-128. doi:10.15366/reice2017.15.1.007
- Hopfenbeck, T.N., Lenkeit, J., El Masri, Y., Cantrell, K., Ryan, J. y Baird, J. (2018). Lessons Learned from PISA: A Systematic Review of Peer- Reviewed Articles on the Programme for International Student Assessment. *Scandinavian Journal of Educational Research*, 62(3), 333-353. doi: 10.1080/00313831.2016.1258726
- Imbernon, F. (2007). Asesorar o dirigir. El papel del asesor/a colaborativo en una formación permanente centrada en el profesorado y en el contexto. *REICE: Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 5(1), 145-152

- Ion, G. y Cano, E. (2012). La formación del profesorado universitario para la implementación de la evaluación por competencias. *Educación XX1*, 15(2), 249-270.
- Ion, G., Silva, P. y Cano, E. (2013). El feedback y el feedforward en la evaluación de las competencias de estudiantes universitarios. *Profesorado. Revista de Curriculum y Formación del Profesorado*, 17(2), 283-301.
- Iranzo, P. (2011). L'avaluació de la formació del professorat al centre: la facilitació per a la transferència. *Recursos i Recerca Educativa de les Illes Balears*, 2, 36-49.
- Jiménez, M., Rodríguez, H., Sánchez, S. y Rodríguez, J. (2018). Construcción del discurso en torno a la Educación Inclusiva. *Revista Electrónica Interuniversitaria de Formación del Profesorado*, 21(1), 185-217.
- Johnson, S., Finlon, K., Kobak, R. y Izard, C. (2017). Promoting Student–Teacher Interactions: Exploring a Peer Coaching Model for Teachers in a Preschool Setting. *Early Childhood Educ*, 45, 461–470.
- Jornet, J.M., González, J., Suárez, J.M. y Peraless, M.J. (2011) Diseño de procesos de evaluación de competencias: consideraciones acerca de los estándares en el dominio de las competencias. *Bordón, Revista de pedagogía*, 63(11), 125-145.
- Kaur, A., Noman, M. y Awang-Hashim, R. (2015). Exploring Strategies of Teaching and Classroom Practices in Response to Challenges of Inclusion in a Thai School: A Case Study. *International Journal of Inclusive Education*, 20(5), 474–485.
- Kaur, A., Noman, M. y Nordin, H. (2017). Inclusive assessment for linguistically diverse learners in higher education. *Assessment & Evaluation in Higher Education*, 42(5), 756-771.

- Kearney, S. (2013). Improving engagement: the use of 'Authentic self-and peer-assessment for learning' to enhance the student learning experience. *Assessment & Evaluation in Higher Education*, 38(7), 875-891.
- Kearney, S., Perkins, T. y Kennedy-Clark, S. (2016). Using self- and peer-assessments for summative purposes: analysing the relative validity of the AASL (Authentic Assessment for Sustainable Learning) model. *Assessment & Evaluation in Higher Education*, 41(6), 840-853.
- Koh, K.H., Tn, C. y Tee, P. (2012) Creating schools through authentic assessment: the case in Singapore. *Educational Assessment, Evaluation and Accountability*, 24, 135–149. doi: 10.1007/s11092-011-9138-y
- Kraft, M. A., Blazar, D. y Hogan, D. (2018). The Effect of Teacher Coaching on Instruction and Achievement: A Meta-Analysis of the Causal Evidence. *Review of Educational Research*, 88(4), 547-588.
- Leahy, S., Lyon, C., Thompson, M., y Wiliam, D. (2005). Classroom assessment that keeps learning on track minute-by-minute, day-by-day. *Educational Leadership*, 63(3), 18-24.
- Levy-Vered, A. y Nasser-Abu, F. (2018). The power of a basic assessment course in changing preservice teachers' conceptions of assessment. *Studies in Educational Evaluation*, 59, 84-93.
- Ley Orgánica 8/2013, de 9 de diciembre, para la mejora de la calidad educativa. Boletín Oficial del Estado, núm. 295, de 10 de diciembre de 2013, pp. 97858 a 97921. Recuperado de <http://www.boe.es/boe/dias/2013/12/10/pdfs/BOE-A-2013-12886.pdf>
- Liesa, E., Castelló, M. y Becerril, L. (2018) Nueva escuela, ¿nuevos aprendizajes?. *Revista de Estudios y Experiencias en Educación*, 2(1), 15-19.
- Liesa, E., Castelló, M., Carretero, M. R., Cano, M. y Mayoral, P. (2012). La atención a la diversidad en la educación secundaria obligatoria:

valoraciones de los profesionales. *Profesorado, Revista de currículum y formación del profesorado*, 16(2), 451-468.

López-Pastor, V. M. (2012). Evaluación Formativa y Compartida en la Universidad: clarificación de conceptos y propuestas de intervención desde la Red Interuniversitaria de Evaluación Formativa. *Psychology, Society & Educación*, 4(1), 113-126.

López-Pastor, V.M. y Pérez-Pueyo, A. (coord.) (2017) *Evaluación formativa y compartida en educación: experiencias de éxito en todas las etapas educativas*. Recuperado de: <https://buleria.unileon.es/handle/10612/5999>

Lukas, J.F. y Santiago, K. (2009). *Evaluación educativa*. Madrid: Alianza Editorial.

Lumsden, G.T. y Michaelides, M.P. (2011). Ecological rationality in teachers' conceptions of assessment across samples from Cyprus and New Zealand. *European Journal of Psychology of Education*, 26, 319–337. doi: 10.1007/s10212-010-0052-3

Maclellan, E. (2004). Authenticity in assessment tasks: a heuristic exploration of academics' perceptions. *Higher Education Research and Development*, 23(1), 19-33.

Marshall, B. y Drummond, M. J. (2006). Cómo se involucran los maestros con la Evaluación para el aprendizaje: lecciones del aula. *Reserach papers in education*, 21(2), 133-149.

Martínez, F. (2013). Dificultades para implementara la evaluación formativa. Revisión de literatura. *Perfiles Educativos*, 139(35), 128-150.

Martínez, J.A. y Cantero, V. (2014). In Search of an Effective Model for Assessing Learning in Bilingual Education: the Authentic Assessment. *Porta Linguarum*, 21(1), 137-150.

Mateo, J. (2000). *La evaluación educativa, su práctica y otras metáforas*. Barcelona, ICE-Horsori.

- Mateo, J. y Martínez, F. (2008). *La evaluación alternativa de los aprendizajes*. Barcelona: Ediciones Octaedro.
- McLaren, S. (2012). Assessment is for learning: supporting feedback. *International Journal of Technology and Design Education*, 22, 227-245.
- Mejías, E. y Monereo, C. (2016). As Life Itself: Authentic Teaching and Evaluation of Professional Consulting Competencies in a Psychology Course. En E. Cano y G. Ion (Eds.) *Innovative Practices for Higher Education Assessment and Measurement* (pp. 311-332) Hershey (Pennsylvania): IGI Global.
- Mejías, E. y Monereo, C. (2018). Transformar avaliação através do Pisa: o Projeto GAPPISA. En M.I. Ortigão (coord.) *Políticas de Avaliação, Currículo e Qualidade: diálogos sobre o Pisa* (pp. 69-90). Brasil: Editora CRV
- Mellati, M. y Khademi, M. (2018). Exploring Teachers' Assessment Literacy: Impact on Learners' Writing Achievements and Implications for Teacher Development. *Australian Journal of Teacher Education*, 43(6). doi: 10.14221/ajte.2018v43n6.1
- Mengual-Andrés, S., Lloret, C. y Roig-Vila, R. (2015). Validación del Cuestionario de evaluación de la calidad de cursos virtuales adaptado a MOOC. *RIED*, 18(2), 145-169.
- Meusen-Beekman, K. D., Joosten-ten Brinke, D., & Boshuizen, H. P. A. (2016). Effects of formative assessments to develop self-regulation among sixth grade students: Results from a randomized controlled intervention. *Studies In Educational Evaluation*, 51, 126-136.
- Meyers, N. y Nulty, D. (2009). How to use (five) curriculum design principles to align authentic learning environments, assessment, students' approaches to thinking and learning outcomes. *Assessment & Evaluation in Higher Education*, 34(5), 565-577.
- Monereo, C. (2003). La evaluación del conocimiento estratégico a través de tareas auténticas. *Revista Pensamiento educativo*, 32, 71-89.

- Monereo, C. (2010). La formación del profesorado: una pauta para el análisis e intervención a través de incidentes críticos. *Revista Iberoamericana de Educación*, 52, 149-178.
- Monereo, C. (2017). The role of critical incidents in the dialogical construction of teacher identity. Analysis of a professional transition case. *Learning, Culture and Social Interaction*, 20, 4-13.
doi: <https://doi.org/10.1016/j.lcsi.2017.10.002>
- Monereo, C. (2019). De l'aprendre per a ser avaluat a l'avaluació per aprendre: perspectives i prospectives. En J. Riera (Ed.). *Reptes de l'educació a Catalunya. Anuari 2018* (pp. 80-127). Barcelona: Fundació Jaume Bofill.
- Monereo, C. (Coord.) (2009). *Pisa como excusa: repensar la evaluación para cambiar la enseñanza*. Barcelona: Graó.
- Monereo, C. (coord.) (2014). *Enseñando a enseñar en la universidad. La formación del profesorado basada en incidentes críticos*. Barcelona: Octaedro - ICE.
- Monereo, C. y Badia, A. (2011). Los heterónimos del docente: Identidad, selfs y enseñanza. C. Monereo y J.I. Pozo. *La Identidad en Psicología de la Educación: necesidad, utilidad y límites* (pp.57-76). Madrid: Narcea.
- Monereo, C. y Castelló, M. (2009). *Las competencias del profesorado para evaluar, y su evaluación*. En C. Monereo. (Coord.), *Pisa como excusa: repensar la evaluación para cambiar la enseñanza* (pp. 71-85). Barcelona: Graó.
- Monereo, C. y Domínguez, C. (2014). La identidad docente de los profesores universitarios competentes. *Educación XX1*, 17(2), 83-104.
- Monereo, C. y Monte, M. (2011). *Docentes en tránsito. Incidentes críticos en secundaria*. Barcelona: Graó.

- Monereo, C. y Morcillo, C. (2011). Las evaluaciones externas como andamiaje de las prácticas de evaluación: un estudio de caso en el área de tecnología. *Cultura y Educación*, 23(2), 183-204.
- Monereo, C. y Pozo, J.I. (2011). *La Identidad en Psicología de la Educación: necesidad, utilidad y límites*. Narcea: Madrid.
- Monereo, C., Badia, A., Bilbao, G., Cerrato, M. y Weise, C. (2009). Ser un docente estratégico: cuando cambiar la estrategia no basta. *Cultura y Educación*, 21(3), 237-256.
- Monereo, C., Castelló, M., Duran, D. y Gómez, I. (2009). Las bases psicoeducativas del proyecto PISA como guía para el cambio en las concepciones y prácticas del profesorado de secundaria. *Infancia y Aprendizaje*, 32(3), 421-447.
- Monereo, C., Panadero, E., y Scartezini, R. (2013). SharEvents. La utilización de informes compartidos sobre incidentes críticos como medio para la formación docente. *Cadernos de Educação*, 42(2), 45-67.
- Monereo, C., Pozo, J.I, y Castelló, M. (2001). *La enseñanza de estrategias de aprendizaje en el contexto escolar*. En Coll. C., Palacios, J. & A. Marchesi (comp.) Desarrollo psicológico y educación, 2. Psicología de la educación escolar (pp. 235-258). Madrid: Alianza
- Monereo, C., Sánchez-Busqués, S. y Suñé, N. (2012). La enseñanza auténtica de competencias profesionales. Un proyecto de aprendizaje recíproco instituto-universidad. *Profesorado, Revista de currículum y formación del profesorado*, 16(1), 79-111.
- Monereo, C., Weise, C. y Álvarez, I. (2013). Cambiar la identidad docente en la universidad. Formación basada en incidentes dramatizados. *Infancia y aprendizaje*, 36(3), 323-340.

- Monforte, G. y Farías, G.M. (2013). La evaluación continua, un incentivo que incrementa la motivación para el aprendizaje. *Revista Iberoamericana de Evaluación Educativa*, 6(2), 265-278.
- Monteagudo, J. y Vera, M.I. (2017). Qué aprenden los alumnos de ESO sobre historia del arte. Un análisis de la legislación y los exámenes en el caso de la Región de Murcia. *Educatio Siglo XXI*, 35(3), 229-254.
- Morales, P. (2009). La evaluación formativa. En P.V. Morales, *Ser profesor: una mirada al alumno* (pp. 41-48). Guatemala: Universidad Rafael Landívar.
- Morales, P. (2010). *Ser profesor: una mirada al alumno*. Guatemala: Universidad Rafael Landívar.
- Moreno, T. (2012). La evaluación de competencias en educación. *Sinéctica*, 39, 1-20.
- Moreno, T. (2015). Las competencias del evaluador educativo. *Revista de la Educación Superior*, 44(2), 101-126.
- Moreno, T. (2017). La evaluación ¿nos conducirá a la tierra prometida?. *Perspectiva Educativa. Formación de Profesores*, 56(1), 147-163.
- Morodo, A. (2017). Identitat discent i disseny universal de l'aprenentatge. Una proposta per a l'atenció de la diversitat. *Revista Catalana de Pedagogia*, 11, 57-77. doi: 10.2436/20.3007.01.86
- Muñoz, A.P., Palacio, M. y Escobar, L. (2011). Teachers' Beliefs About Assessment in an EFL Context in Colombia. *PROFILE*, 14(1), 143-158.
- Muñoz, J. y Charro, E. (2017). Los ítems PISA como herramienta para el docente en la identificación de los conocimientos y habilidades científicas. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 14(2), 317-338.
- Nail, O., Gajardo, J. y Muñoz, R. (2012). La técnica de análisis de incidentes críticos: una herramienta para la reflexión sobre prácticas docentes en convivencia escolar. *Psicoperspectivas*, 11(2), 56-76.

- OECD (2010). La naturaleza de l'aprenentatge. Utilitzar la recerca per inspirar la pràctica. Recuperado de:
https://www.fbofill.cat/sites/default/files/The_Nature_of_Learning-Practitioner_Guide-CAT%5B3%5D%20Copy.pdf
- Oller, M., Navas, C., y Carrera, J. (2018). Docència compartida a l'aula: Reptes i possibilitats. *Guix*, 448, 43-48.
- Opazo, M., Pérez-Cabaní, M.L. y Sepúlveda, A. (2016). Del dicho al hecho en la evaluación del aprendizaje en la universidad: la conveniencia de evaluar a través de tareas auténticas. *Contextos*, 36, 127-143.
- Opazo, M., Sepúlveda, A. y Pérez-Cabaní, M.L. (2015). Estrategias de evaluación del aprendizaje en la universidad y tareas auténticas: percepción de los estudiantes. *Diálogos educativos*, 29(1), 19-3.
- Opre, D. (2015). Teachers' Conceptions of Assessment. *Procedia - Social and Behavioral Sciences*, 209, 229 – 233. doi:10.1016/j.sbspro.2015.11.222
- Orden EDU/849/2010, de 18 de marzo, por la que se regula la ordenación de la educación del alumnado con necesidad de apoyo educativo y se regulan los servicios de orientación educativa en el ámbito de gestión del Ministerio de Educación, núm. 83, de 6 de abril de 2010, pp. 31332 a 31380.
Recuperado de <https://www.boe.es/boe/dias/2010/04/06/pdfs/BOE-A-2010-5493.pdf>
- Panadero, E. y Alonso-Tapia, J. (2013). Autoevaluación: connotaciones teóricas y prácticas. cuándo ocurre, cómo se adquiere y qué hacer para potenciarla en nuestro alumnado. *Electronic Journal of Research in Educational Psychology*, 30(11), 551-576.
- Panadero, E., Andrade, H., y Brookhart, S. M. (2018). Fusing self-regulated learning and formative assessment: A roadmap of where we are, how we got here, and where we are going. *The Australian Educational Researcher*, 45(1), 13-31.

- Pascual-Gómez, I., Lorenzo-Llamas, E.M. y Monge-López, C. (2015). Análisis de validez en la evaluación entre iguales: un estudio en educación superior. *RELIEVE*, 21(1). doi: 10.7203/relieve.21.1.5168
- Pérez-Cabaní, M.L. y Carretero, M.R. (2009). La evaluación del aprendizaje en la educación secundaria. Análisis de un proceso de cambio. *Límite: revista de filosofía y psicología*, 19, 93-126.
- Pérez-Cabaní, M.L., Juandó, J. y Argelagós, E. (2015) Analysing a management model for competence-based university teaching: the lecturer's perspective / Análisis de un modelo de gestión de la docencia basado en competencias en la universidad: la perspectiva del profesor. *Infancia y Aprendizaje*, 38(4), 872-908. doi: 10.1080/02103702.2015.1076270
- Perrenoud, P. (2004). *Diez nuevas competencias para enseñar*. Barcelona: Graó.
- Popham, W.J. (2004). Why assessment illiteracy is professional suicide. *Educational Leadership*, 62(1), 82-83.
- Popham, W.J. (2009). Assessment Literacy for Teachers: Faddish or Fundamental?. *Theory Into Practice*, 48(1), 4–11.
- Postareff, L., Virtanen, V., Katajavuori, N. y Lindblom-Ylänne, S. (2012). Academics' conceptions of assessment and their assessment practices. *Studies in Educational Evaluation*, 38(3-4), 84-92.
- Powell, D. R., Diamond, K. E., Burchinal, M. R. y Koehler, M. J. (2010). Effects of an Early Literacy Professional Development Intervention on Head Start Teachers and Children. *Journal of Educational Psychology*, 102(2), 299-312.
- Ramírez, E. (2016). Evaluación de las competencias básicas en primaria: una mirada desde la óptica docente. *Profesorado. Revista de currículum y formación del profesorado*, 20(1), 243-264.
- Raymond, J. E., Homer, C. S. E., Smith, R. y Gray, J. E. (2013). Nurse Education in Practice Learning through authentic assessment: An evaluation of a new

development in the undergraduate midwifery curriculum. *Nurse Education in Practice*, 13(5), 471–476.

Remesal, A. (2011). Primary and secondary teachers' conceptions of assessment: A qualitative study. *Teaching and Teacher Education*, 27(2), 472–482.

Robinson, V. M., Hohepa, M. K. y Lloyd, C. (2009). Leading the improvement of teaching and learning. Dentro de *School Leadership and Student Outcomes: Identifying What Works and Why. Best Evidence Synthesis Iteration (BES)*. Wellington, Nova Zelanda: Ministry of Education.

Rodríguez, G. y Ibarra, M. S. (2012). Reflexiones en torno a la competencia evaluadora del profesorado en la Educación Superior. *Revista de Docencia Universitaria*, 10(2), 149-161.

Rodríguez, M. E. (2010). Exploraciones en torno a concepciones de docentes sobre la evaluación. *Enunciación*, 5(1), 8-17.

Roig-Vila, R. y Rosales-Statkus, S. (2016). El relato digital. Análisis de sus elementos y tipología. *RIITE. Revista Interuniversitaria de Investigación en Tecnología Educativa*, 0, 84-94. Doi: <http://dx.doi.org/10.6018/riite/2016/257951>

Romero-Martín, R., Castejón-Oliva, F., López-Pastor, V. y Fraile-Aranda, A. (2017). Formative assessment, communication skills and ICT in Initial Teacher Education. *Comunicar*, 52(3), 73-82. doi: <https://doi.org/10.3916/C52-2017-07>

Rueda, M. y Torquemada, A. (2008). Las concepciones sobre "evaluación" de profesores y estudiantes: sus repercusiones en la evaluación del desempeño docente. *REencuentro. Análisis de Problemas Universitarios*, 53(4), 97-112.

Ruiz-Terroba, R., Vázquez, E. y Sevillano-García, M.L. (2017). La rúbrica de evaluación de la competencia en expresión escrita. Percepción del

alumnado sobre su funcionalidad. *Ocnos, Revista de Estudios sobre lectura*, 16(2), 107-117.

Sabirón, F. y Arraiz, A. (2013) Aprendiendo de la evaluación: decálogo para la evaluación auténtica de competencias profesionales a través del portafolio. *Revista Iberoamericana de Evaluación Educativa*, 6(1), 135-152.

Sala, I., Sánchez, S., Giné, C. y Díez, E. (2013). Análisis de los distintos enfoques del paradigma del diseño universal aplicado a la educación. *Revista Latinoamericana de Educación Inclusiva*, 8(1), 143-152.

Salgado, J. y Hermans, H. (2005). The return of subjectivity: From a multiplicity of selves to the dialogical self. *E-journal of Applied Psychology*, 1(1), 3-13.

Salinas, D. (2002). *¡Mañana examen! La evaluación: entre la teoría y la realidad*. Barcelona: Graó.

Sanahuja, A. y Sánchez-Tarazaga, L. (2018). La competencia evaluativa de los docentes: formación, dominio y puesta en práctica en el aula. *Revista Iberoamericana de Educación*, 76(2), 95-116.

Sanmartí, N. (2007). *Evaluar para aprender*. Barcelona: Graó.

Sanmartí, N. (2010a). Aprender a evaluarse: motor de todo aprendizaje. *Aula de Innovación Educativa*, 192, 26-29.

Sanmartí, N. (2010b). Avaluar per aprendre. L'avaluació per millorar els aprenentatges de l'alumnat en el marc del currículum per competències. Generalitat de Catalunya. Recuperado de:
http://xtec.gencat.cat/web/.content/alfresco/d/d/workspace/SpacesStore/0024/fc53024f-626e-423b-877a-932148c56075/avaluar_per_aprendre.pdf

Sanmartí, N. y Sardà, A. (2007). Luces y sombras en la evaluación de competencias: el caso de PISA. *Cuadernos de Pedagogía*, 370, 60-63.

Santos-Guerra, M. A. (2003). *Una Flecha en la diana: la evaluación como Aprendizaje*. Madrid Narcea.

- Segers, M. y Tillema, H. (2011). How do Dutch secondary teachers and students conceive the purpose of assessment?. *Studies in Educational Evaluation*, 37, 49–54. doi: 10.1016/j.stueduc.2011.03.008
- Smith, K. L., Rodgers, M., Hawk, E. y Fulcher, K.H. (2015) Communication is Key: Unpacking “Use of Assessment Results to Improve Student Learning”. *Research & Practice in Assessment*, 10, 15-29.
- Sondergeld, T. A., Bell. C. A. y Leusner, D. M. (2010). Understanding how Teachers Engage in Formative Assessment. *Teaching & Learning*, 24(2), 72-86.
- Stiggins, R. (2002). Learning teams for assessment literacy. *Journal of Staff Development*, 30(4), 5-7.
- Swaffield, S. (2011). Getting to the heart of authentic Assessment for Learning. *Assessment in Education: Principles, Policy & Practice*, 18(4), 433-449.
- Syrene, S. A., Venville, G. y Chapman, A. (2009). Classroom Assessment: Juxtaposing Teachers’ Beliefs with Classroom Practices . Australian Association For Research In Education: International Education Research Conference (29 November- 3 December 2009)
- Taras, M. (2015). Autoevaluación del estudiante: ¿Qué hemos aprendido y cuáles son los desafíos?. *RELIEVE*, 21(1). doi: 10.7203/relieve.21.1.6394
- Tierney, R.D. (2006). Changing practices: influences on classroom assessment. *Assessment in Education*, 13(3), 239–264.
- Tillema, H., Leenknecht, M. y Segers, M. (2011). Assessing assessment quality: Criteria for quality assurance in design of (peer) assessment for learning - A review of research studies. *Studies in Educational Evaluation*, 37, 25-34.
- Tonda, P. y Medina, A. (2013). La formación del profesorado en la competencia evaluadora: un camino hacia la calidad educativa. *Enseñanza & Teaching*, 31(2), 167-188. Recuperado de: <http://revistas.usal.es/index.php/0212-5374/article/view/11632>

- Trevitt, C. y Stocks, C. (2012). Signifying authenticity in academic practice: a framework for better understanding and harnessing portfolio assessment. *Assessment & Evaluation in Higher Education*, 37(2), 245-257.
- Tufail, M., Murtaza, S. y Iqbal, R. (2016) Bloom's Taxonomy: Improving Assessment and Teaching-Learning Process. *Journal of Education and Educational Development*, 3(2), 203-221.
- Vallejo, M. y Molina, J. (2014). La evaluación auténtica de los procesos educativos. *Revista Ibero-Americana de Educação*, 64, 11-25.
- Van Venn, K. y Lasky, S. (2005). Emotions as a lens to explore teacher identity and change: Different theoretical approaches. *Teaching and Teacher Education*, 21(8), 895-898.
- Vergara, C. E. (2011). Concepciones de Evaluación del Aprendizaje de Docentes Chilenos Destacados de Educación Básica. *Acción Pedagógica*, 20, 6-18.
- Villarroel, V., Bloxham, S., Bruna, D., Bruna, C. y Herrera-Seda, C. (2017). Authentic assessment: creating a blueprint for course design. *Assessment & Evaluation in Higher Education*, 43(5), 840-854.
- Villarroel, V., Boud, D., Bloxham, S. Bruna, D. y Bruna, C. (2019). Using principles of authentic assessment to redesign written examinations and tests. *Innovations in Education and Teaching International*. doi: 10.1080/14703297.2018.1564882
- Villarroel, V., Bruna, D., Bustos, C., Bruna, C. y Márquez, C. (2018). Análisis de pruebas escritas bajo los principios de la evaluación auténtica. Estudio comparativo entre carreras de la salud y otras carreras de dos universidades de la Región del Biobío. *Revista Médica de Chile*, 146, 45-52.
- Voerman, L., Meijer, P., Korthagen, F. y Simons, R. (2012). Types and frequencies of feedback interventions in classroom interaction in secondary

education. *Teaching and Teacher Education*, 28, 1107-1115. doi:
10.1016/j.tate.2012.06.006

Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge: Harvard University Press.

Weise, C. y Sánchez-Busqués, S. (2013). Identidad docente y estrategias de resolución de incidentes críticos en contextos universitarios de alta diversidad sociocultural. *Cultura y Educación*, 25(4), 561-576.

Wiggins, G. (2011). Giving Students a Voice: The power of Feedback to Improve Teaching. *Educational Horizons*, 89(3), 23-26.

William, D., Lee, C., Harrison, C. y Black, P. (2004). Teachers developing assessment for learning: impact on student achievement. *Assessment in Education*, 11(1), 49-65.

Xarxa de Competències Bàsiques (2019). Formar per transformar. La Xarxa de Competències Bàsiques, un model de professionalització docent.

Recuperado de:

<http://xtec.gencat.cat/web/.content/curriculum/xarxacb/documents/El-model-formatiu-XCB-copy.pdf>

Xie, K., Kyu, M., Sheng-Lun, C. y Luthy, N. (2017). Teacher professional development through digital content evaluation. *Education Technology, Research and Development*, 65(4), 1067–1103.

Yamila, D. (2016). Autorregulación y rúbricas como herramienta de evaluación. Experiencia desarrollada en educación primaria. *Escuela Abierta*, 19, 65-79.

ANEXOS



Guía GAPpisa

La Guía GAPPisa es un instrumento de auto-reflexión para analizar los niveles de competencialidad, inclusividad y poder formativo de las pruebas de evaluación.

➤ ¿Qué entendemos por competencialidad?

Una prueba de evaluación de competencias nos tiene que permitir evaluar el grado en que un alumno es capaz de **coordinar diferentes tipos de conocimiento (conceptos, procedimientos, actitudes y estrategias) para solucionar un problema prototípico del ámbito** que evaluaremos. Por lo tanto, no hay suficiente con evaluar solo conceptos o procedimientos de forma aislada

➤ ¿Qué entendemos por inclusividad?

Una evaluación inclusiva es aquella que se **ajusta a la diversidad de perfiles de todo el grupo clase**. Para conseguirlo tenemos que ajustar, no solo la complejidad de las tareas, sino también el formato de los enunciados y de las respuestas. Todo alumno tiene derecho a recibir información en un formato que se ajuste a sus preferencias y a defender su aprendizaje en una modalidad de respuesta y condiciones que le permitan reforzar sus fortalezas y compensar sus debilidades.

➤ ¿Qué entendemos por poder formativo?

La evaluación formativa y formadora es aquella que **permite a alumnos y profesores aprender gracias a la realización de la prueba**. Para conseguirlo será necesario que antes, durante y después de la prueba, guiemos a los alumnos en: a) la interiorización de los criterios de calidad; b) la mejor gestión de los recursos disponibles; c) la construcción de nuevo aprendizaje; i d) la conceptualización del error como oportunidad de aprendizaje. Paralelamente, los docentes aprenderemos: a) qué necesidades tienen los alumnos; b) cómo diseñar las próximas ayudas educativas; i c) cómo estructurar y organizar la docencia en próximas ocasiones. En este marco, no existen diferencias entre una actividad de aprendizaje y una de evaluación.

Estructura y uso de la guía

La guía está estructurada en 10 dimensiones que hacen referencia a diferentes aspectos de una prueba o actividad de evaluación. Los indicadores de calidad de cada dimensión se plantean a modo de pregunta para que puedas valorar cuál es la situación actual de tu prueba y, si es necesario, qué necesitas añadir o modificar para mejorarla.

Para utilizar la guía te recomendamos que primero hagas una lectura general y te asegures de entender cada aspecto. Si es necesario puedes consultar el material didáctico específico para cada dimensión. Una vez te hayas familiarizado con las diferentes preguntas, selecciona la prueba de evaluación que quieras analizar y responde a cada cuestión en el apartado Reflexión. Anota todas las mejoras que te gustaría introducir en la prueba y, finalmente, rediseña la "versión mejorada" de tu actividad de evaluación.



Guía GAPPISA de SINTEdi está sujeta a una licencia de Reconocimiento-NoComercial 4.0 Internacional de Creative Commons

Guía GAPPISA

Una evaluación de competencias, formativa e inclusiva

	Dimensión	Indicadores de calidad	Reflexión
1	Situación problema inicial auténtica	<p>¿La prueba plantea una situación problema inicial suficientemente contextualizada?</p> <p><i>*Necesario para evaluar competencias</i></p>	
2	Actividades auténticas	<p>¿Las actividades de la prueba son realistas?</p> <p><i>*Necesario para evaluar competencias</i></p> <p>¿Las actividades de la prueba son relevantes?</p> <p><i>*Necesario para evaluar competencias</i></p> <p>¿Las actividades de la prueba son socializadoras?</p> <p><i>*Necesario para evaluar competencias</i></p>	
3	Actividades de diferente complejidad cognitiva y regulación emocional	<p>¿Existen demandas de baja, media y elevada complejidad cognitiva?</p> <p><i>*Necesario para evaluar competencias</i></p> <p>¿Los niveles de complejidad son progresivamente más elevados (modelaje)?</p> <p><i>*Necesario para realizar una evaluación formativa</i></p> <p>¿Se ajusta la complejidad de la prueba a las necesidades del alumno pero sin dejar de evaluar competencias?</p> <p><i>*Necesario para realizar una evaluación inclusiva</i></p>	
4	Formato heterogéneo de los enunciados	<p>¿Las preguntas se presentan en diversidad de formatos (in-put)? por ejemplo: texto, imagen, vídeo, gráfico, etc.</p> <p><i>*Necesario para evaluar competencias</i></p> <p>¿Se ajusta el formato de los enunciados a las necesidades del alumno?</p> <p><i>*Necesario para realizar una evaluación inclusiva</i></p>	
5	Formato heterogéneo de las respuestas	<p>¿La prueba requiere del alumno diversas modalidades de respuesta? Por ejemplo: escrita, oral, gráfica, etc.</p> <p><i>*Necesario para evaluar competencias</i></p>	



Guía GAPPISA de SINTEdi está sujeta a una licencia de Reconocimiento-NoComercial 4.0 Internacional de Creative Commons

		<p>¿Se ajusta el formato de las respuestas a las necesidades del alumno?</p> <p><i>*Necesario para realizar una evaluación inclusiva</i></p>	
6	Condiciones de resolución de la prueba	<p>¿Las condiciones de resolución de la prueba son similares a las de la vida real? Por ejemplo: acceso a ayudas y recursos, trabajo en equipo, etc.</p> <p><i>*Necesario para evaluar competencias</i></p> <p>¿Se ajustan las condiciones de resolución a las necesidades de los alumnos?</p> <p><i>*Necesario para realizar una evaluación inclusiva</i></p>	
7	Evaluación del proceso	<p>¿La prueba requiere que el alumno explicita el proceso de resolución en algún momento (antes/durante/después)?</p> <p><i>*Necesario para realizar una evaluación formativa</i></p> <p>¿Se evalúa el ajuste del alumno a situaciones nuevas /inesperadas?</p> <p><i>*Necesario para evaluar competencias</i></p> <p>¿Se evalúa el uso que el alumno hace de las ayudas/recursos?</p> <p><i>*Necesario para realizar una evaluación formativa</i></p>	
8	Puntuaciones y criterios compartidos	<p>¿Se comparte con el alumno el valor de cada pregunta?</p> <p><i>*Necesario para realizar una evaluación formativa</i></p> <p>¿Se comparten con el alumno los criterios de corrección cualitativos?</p> <p><i>*Necesario para realizar una evaluación formativa</i></p> <p>¿Se fomenta de forma eficiente la interiorización de los criterios de calidad?</p> <p><i>*Necesario para realizar una evaluación formativa</i></p>	

9	Comunicación de los resultados	Comunicación bidireccional, cualitativa y personalizada.
		¿Se contextualizan los resultados del alumno con su rendimiento en pruebas anteriores (progreso personal)?
		<i>*Necesario para realizar una evaluación formativa</i>
		¿El tipo de correcciones y evidencias que se ofrecen al alumno son suficientes para fomentar su aprendizaje? ¿Hacen referencia a la eficiencia del proceso de resolución y corrección de los resultados?
		<i>*Necesario para realizar una evaluación formativa</i>
		¿Se ofrece la posibilidad de revisar la propia prueba y volverla a realizar de forma que se pueda subir nota?
		<i>*Necesario para realizar una evaluación formativa</i>
		¿Se ajusta la corrección y la comunicación de la prueba a las necesidades del alumno?
		<i>*Necesario para realizar una evaluación formativa</i>
10	Atención a la diversidad	¿Se han analizado los perfiles de aprendizaje de todo el grupo clase?
		<i>*Necesario para realizar una evaluación inclusiva</i>
		¿Se han analizado los intereses y motivaciones de los alumnos?
		<i>*Necesario para realizar una evaluación inclusiva</i>

Otras observaciones:

Anexo 2. Rúbrica de evaluación de las pruebas autenticadas

Hacia una evaluación competencial, formativa e inclusiva.

Criterios de calidad para elaborar pruebas de evaluación válidas, fiables y justas.

Pseudo-mejoras		Bien, estás mejorando tu prueba			
-0.5	Dimensión	0.25 puntos	0.75 puntos	1 puntos	Extra
Ejemplo: "Estás en una entrevista de trabajo y te piden que realices este examen"	1. Situación problema inicial realista	Evidencia real. Se empieza la prueba con un texto o problema real, pero se continua con una demanda no realista. <i>*Evaluación competencial</i>	Demanda real. La prueba plantea una situación problema inicial realista (que incluye una demanda realista). <i>*Evaluación competencial</i>	Contexto extraacadémico real. Se describen detalladamente las características del contexto físico y social en el que se enmarca esta situación problema (actores, objetivos, audiencia, etc.) <i>*Evaluación competencial</i>	
-0.5	Dimensión	0.25 puntos	0.75 puntos	1 puntos	Extra
Ejemplo: Rellenar la prueba con fotografías e imágenes que poco o nada tienen que ver con la demanda...	2. Actividades auténticas	Las actividades de la prueba son realistas . <i>*Evaluación competencial</i>	Las actividades de la prueba son realistas y relevantes . <i>*Evaluación competencial</i>	Las actividades de la prueba son realistas, relevantes y socializadoras (fomentan las habilidades de interacción social). <i>*Evaluación competencial</i>	
-0.5	Dimensión	0.25 puntos	0.75 puntos	1 puntos	Extra
Ejemplo: Disminuir mucho el nivel de la prueba (cursos inferiores) o a la inversa.	3. Actividades de diferente complejidad cognitiva y emocional	¿Existen demanda de baja, media y alta complejidad cognitiva/emocional? <i>*Evaluación competencial</i>	¿Los niveles de complejidad son progresivamente más elevados (modelaje)? <i>*Evaluación formativa</i>	¿Se ajusta la complejidad de la prueba a las necesidades de los alumnos (con NESE) pero sin dejar de evaluar competencias?	

				<i>*Evaluación inclusiva</i>	
-0.5	Dimensión	0.25 puntos	0.75 puntos	1 puntos	Extra
Ejemplo: Cambiar el tipo fuente de letra de los enunciados.	4. Formato heterogéneo de los enunciados	¿Los enunciados (in-put) se presentan en un mínimo de 2 formatos básicos distintos? - P. e.: texto (diferentes tipos de texto) + imagen/gráfico <i>*Evaluación competencial</i>	Los enunciados (in-put) se presentan en un mínimo de 3 formatos básicos distintos? - P. e.: texto (diferentes tipos de texto) + imagen + audiovisual <i>*Evaluación competencial</i>	¿Se ajusta el formato de los enunciados a las necesidades especiales de los alumnos que lo requieran? <i>*Evaluación inclusiva</i>	
-0.5	Dimensión	0.25 puntos	0.75 puntos	1 puntos	Extra
Ejemplo: Preguntas tipo test, de relacionar, de rellenar espacios en blanco, etc. (todas ellas escritas y textuales).	5. Formato heterogéneo de las respuestas	La prueba requiere del alumno al menos dos modalidades de respuesta diferentes? P. E.: textual + oral. <i>*Evaluación competencial</i>	La prueba requiere del alumno al menos tres modalidades de respuesta diferentes? P. E.: textual + oral + gráfica. <i>*Evaluación competencial</i>	Se ajusta el formato de las respuestas a las preferencias del alumno? <i>*Evaluación inclusiva</i>	
-0.5	Dimensión	0.25 puntos	0.75 puntos	1 puntos	Extra
Ejemplo: Los alumnos pueden realizar un máximo de 3 preguntas al "experto" (profesor). Se les bajará 0.25 puntos por pregunta realizada.	6. Condiciones de resolución de la prueba	Las condiciones de resolución de la prueba son similares a las de la vida real, con relación a acceso a recursos y ayudas . <i>*Evaluación competencial</i>	Las condiciones de resolución de la prueba son similares a las de la vida real, en relación (también) a la ubicación, temporalización, consulta a expertos, interacción social , etc. <i>*Evaluación competencial</i>	¿Se ajustan las condiciones de resolución a las necesidades de los alumnos? <i>*Evaluación inclusiva</i>	
-0.5	Dimensión	0.25 puntos	0.75 puntos	1 puntos	Extra
Ejemplo: Pregunta de baja complejidad (p. e. relacionar	7. Evaluación del proceso	¿Se evalúa el ajuste del alumno a situaciones nuevas/inesperadas?	¿La prueba requiere que el alumno explícite el proceso de	¿Se evalúa el uso que el alumno hace de las	

vocabulario inglés-castellano), con una única respuesta correcta, que implica reproducir conocimiento.		<i>*Evaluación competencial</i>	resolución en algún momento (antes/durante/después)? <i>*Evaluación formativa</i>	ayudas/recursos y aprovechamiento de las condiciones? <i>*Evaluación formativa</i>	
-0.5	Dimensión	0.25 puntos	0.75 puntos	1 puntos	Extra
Ejemplo: "Cada pregunta vale 2 puntos". Rúbrica: <i>Mapa de conceptos:</i> 0. <i>No lo hace o lo hace mal.</i> 1. <i>Lo hace regular.</i> 2. <i>Lo hace correctamente.</i>	8. Puntuaciones y criterios compartidos	¿Se comparte con el alumno el valor de cada pregunta? <i>*Evaluación formativa</i>	¿Se comparten con el alumno los criterios de corrección cualitativos? <i>*Evaluación formativa</i>	¿Se fomenta de forma eficiente la interiorización de los criterios de corrección/calidad cualitativos? (momento=antes, durante y después; instrumento=rúbrica; etc.). <i>*Evaluación formativa</i>	
-0.5	Dimensión	0.25 puntos	0.75 puntos	1 puntos	Extra
Ejemplo: "Hacemos correcciones cualitativas, revisamos la prueba en clase, preparamos tutorías individuales, etc." ¡Pero el examen siempre se debe quedar en el cole y no se lo pueden llevar para preparar el siguiente!	9. Comunicación de los resultados	- Comunicación bidireccional y cualitativa ? - Se contextualizan los resultados del alumno con su rendimiento en pruebas anteriores (progreso personal)? <i>*Evaluación formativa</i>	- El tipo de correcciones y evidencias que se ofrecen al alumno son suficientes para fomentar su aprendizaje y hacen referencia tanto al proceso como al resultado? <i>*Evaluación formativa</i>	- Se ofrece la posibilidad de revisar la prueba (una vez corregida) y volverla a realizar para mejorar la nota? - Se ajusta la corrección y comunicación de los resultados de la prueba a las necesidades del alumno? <i>*Evaluación inclusiva</i>	
-0.5	Dimensión	0.25 puntos	0.75 puntos	1 puntos	Extra

<p>Ejemplo: Dar a todos los alumnos el mismo examen y a algunos decirles “no hagáis las preguntas más difíciles, que son la 5, 6 y 7” (encima, tacharles estas preguntas delante de todos).</p>	<p>10. Atención a la diversidad</p>	<p>¿Se han analizado los puntos fuertes y débiles de los alumnos? *Evaluación inclusiva</p>	<p>Se han analizado los distintos perfiles de aprendizaje del grupo clase considerando: preferencias de in-put y out-put; funciones ejecutivas; tipo de interacción; etc.? *Evaluación inclusiva</p>	<p>¿Se han analizado los diferentes perfiles de aprendizaje y su perfil motivacional? *Evaluación inclusiva</p>	
<p>Puntuación total</p>		<p>Observaciones</p>			
		<p>Equipo Gappisa Restamos a tu disposición para cualquier comentario, duda o reflexión en proyectogappisa@gmail.com</p>			

Anexo 3. Informe Final sobre la aplicación de la prueba

Proyecto GAPPISA



ACTIVIDAD FORMATIVA 5

Aplicación de la prueba de evaluación y recolección de evidencias

El siguiente informe se ha elaborado a modo de entrevista escrita y consta de 12 preguntas. Como verás, hace referencia a las percepciones y evidencias recolectadas en los tres momentos del proceso de aplicación de la prueba: antes, durante y después. Por favor, responde a cada pregunta en detalle y profundidad. Tiempo estimado de respuesta del informe: 30 – 45 minutos.

INFORME FINAL: Percepciones y evidencias

Nombre del docente	
Centro educativo	
Años de experiencia docente	
Unidad didáctica evaluada	
Área/ materia	
Nivel educativo	

*Recuerda, el informe está organizado en tres secciones en relación a la aplicación de la prueba: **ANTES**, **DURANTE** y **DESPUÉS**. Te recomendamos que hagas una lectura en diagonal de todo el informe antes de empezar a responder las preguntas para así poder distribuir la información correctamente.

ANTES DE LA APLICACIÓN DE LA PRUEBA

1.1 ¿Has realizado cambios en relación al diseño / planificación habitual de la enseñanza? Si la respuesta es afirmativa, ¿qué aspectos has cambiado y, brevemente, en qué han consistido?

1.2 ¿Has realizado cambios en relación a la enseñanza habitual en el aula? Si la respuesta es afirmativa, ¿qué aspectos has cambiado y, brevemente, en qué han consistido?

1.3 ¿Has realizado cambios en relación a las orientaciones que das habitualmente para la preparación de la prueba por parte de los alumnos? Si la respuesta es afirmativa, ¿qué aspectos has cambiado y, brevemente, en qué han consistido?

1.4 ¿Has percibido cambios en relación a la forma de aprender y/o estudiar habitual de los alumnos, o de algún alumno o grupo de alumnos en particular? Si la respuesta es afirmativa, ¿qué aspectos han cambiado y, brevemente, en qué han consistido?

***¿Qué variables has contemplado y qué evidencias has recolectado que te hayan permitido llegar a la afirmación anterior? ¿Cómo lo has hecho?** *Por ejemplo: he elaborado un breve cuestionario sobre su proceso de estudio que han rellenado al finalizar la prueba; les he preguntado oralmente en clase; realizamos una sesión de discusión / valoración de la prueba; etc. Y las variables que he contemplado han sido: cantidad de estudio, tipo de estrategias (memorización, comprensión, organización de los materiales, etc.), finalidad, etc.*

DURANTE LA APLICACIÓN DE LA PRUEBA

2.1 ¿Has percibido cambios en relación a la forma habitual en que presentas o explicas qué debe hacerse en las pruebas de evaluación? Si la respuesta es afirmativa, ¿qué aspectos has cambiado y, brevemente, en qué han consistido?

2.2 ¿Has percibido cambios en relación a la forma habitual en que los alumnos realizan las pruebas de evaluación? Si la respuesta es afirmativa, ¿qué aspectos han cambiado y, brevemente, en qué han consistido?

**¿Qué variables has contemplado y qué evidencias has recolectado que te hayan permitido llegar a la afirmación anterior? ¿Cómo lo has hecho? Por ejemplo: he analizado sus respuestas; les he preguntado oralmente sobre el proceso que han seguido; he entrevistado a 3 o 4 alumnos con rendimiento heterogéneo (muy bueno, bueno, regular, bajo); etc. Las variables que he contemplado han sido: grado de planificación y revisión de sus respuestas, orden, etc.*

2.3 ¿Ha habido alumnos que han presentado resistencias a realizar dicha prueba de evaluación? Si la respuesta es afirmativa, ¿qué tipo de resistencias?

**¿Qué variables has contemplado o qué evidencias has recolectado que te hayan permitido llegar a la afirmación anterior? Por ejemplo: algunos alumnos me han comentado abiertamente que...; en la sesión de revisión en clase, algunos se quejaron de...; etc.*

DESPUÉS DE LA APLICACIÓN DE LA PRUEBA

3.1 ¿Has realizado cambios en relación con la revisión/corrección habitual de una prueba de evaluación? Si la respuesta es afirmativa, ¿qué aspectos has cambiado y, brevemente, en qué han consistido?

3.2 ¿Has realizado cambios en relación a la forma habitual en que comunicas a los alumnos los resultados de la evaluación? Si la respuesta es afirmativa, ¿qué aspectos has cambiado y, brevemente, en qué han consistido?

3.3 ¿Has percibido cambios en relación a la forma habitual en que los alumnos reciben los resultados de una prueba de evaluación? Si la respuesta es afirmativa, ¿qué aspectos han cambiado de la forma habitual de comportarse de los alumnos y, brevemente, en qué han consistido?

***¿Qué variables has contemplado y qué evidencias has recolectado que te hayan permitido llegar a la afirmación anterior? ¿Cómo lo has hecho?** *Por ejemplo: he observado sus reacciones, he realizado tutorías, les he preguntado directamente, etc. Las variables que he contemplado han sido: grado de satisfacción con el resultado, habilidad del alumno para anticipar la calificación que recibiría (ajuste a sus expectativas), nivel de motivación para mejorar y superarse, etc.*

3.4 ¿Has percibido cambios en relación a los resultados habituales de aprendizaje de los alumnos (calificaciones)? Si la respuesta es afirmativa, ¿qué aspectos han cambiado de las calificaciones, y, brevemente, en qué han consistido?

**¿Qué variables has contemplado o qué evidencias has recolectado que te hayan permitido llegar a la afirmación anterior? ¿Cómo lo has hecho? Por ejemplo: he comparado la calificación de cada alumno con la media que tenía de las evaluaciones anteriores, etc. Las variables que he contemplado han sido: porcentaje de alumnos que han superado la prueba, nivel en que el rendimiento de determinados alumnos se ha visto afectado (p. e. alumnos que antes sacaban notas medias y ahora han aumentado su rendimiento y viceversa), etc.*

3.5 ¿Has percibido cambios en relación al grado de satisfacción de los alumnos? ¿Qué aspectos les han gustado más /menos de la prueba de evaluación auténtica en comparación con las pruebas de evaluación a las que estaban acostumbrados anteriormente?

**¿Qué variables has contemplado y qué evidencias has recolectado que te hayan permitido llegar a la afirmación anterior? ¿Cómo lo has hecho? Por ejemplo: he elaborado un breve cuestionario que me ha permitido recoger sus percepciones, motivación y nivel de satisfacción, hemos realizado una discusión en clase sobre las siguientes variables, etc. Las variables que he contemplado han sido: nivel de dificultad percibida (p. e. ¿esta prueba de evaluación os ha parecido más fácil o difícil respecto a las que veníamos haciendo?); utilidad (p. e. ¿qué tipo de evaluación o examen os parece más útil? ¿Por qué?); relevancia (p. e. ¿los contenidos que hemos evaluado / aplicado en esta prueba son relevantes en vuestro día a día? Y ¿para vuestro futuro?); motivación (p. e. ¿estabais más motivados para realizar esta prueba? ¿Por qué creéis que ha pasado?); valoración (p. e. ¿qué tipo de prueba te gusta más: ésta o las anteriores? ¿Por qué?); etc.*

Anexo 4. Análisis de los estudios sobre concepciones de los docentes sobre evaluación

	Por qué se evalúa	Qué se evalúa	Quiénes participan en la evaluación	Cómo es el proceso y/o de sistema de evaluación	Con qué se evalúa	Cómo se comunican y para qué se utilizan los resultados	Cómo es la cultura de centro sobre evaluación
MacIellan (2004) Universidad. Varias disciplinas.	Respuestas abiertas sobre propósitos.	Respuestas abiertas sobre el tipo de aprendizaje a evaluar y modos de usarlo.	Respuestas abiertas sobre agentes que deben participar en la evaluación formativa y en la sumativa.	Respuestas abiertas sobre el uso y conocimiento de los criterios de evaluación y los estándares de aprendizaje esperados.		Respuestas abiertas sobre la importancia de la retroalimentación.	
Rueda y Torquemada (2008) Universidad. Varias disciplinas.	Respuestas abiertas, sobre propósitos: - medición y retención del conocimiento - fines acreditativos	Respuestas abiertas: - actitudes - rendimiento del alumno - cumplimiento de objetivos		Respuestas abiertas sobre el proceso: - subjetivo y ambiguo - es normativo e incompleto - continuo e integral	Respuestas abiertas: - exámenes - variedad de instrumentos de carácter práctico - criterios definidos		Respuestas abiertas: - aspectos negativos - aspectos positivos
Abdul et al. (2009) Universidad. Matemáticas.	Respuestas abiertas sobre propósitos, agrupadas en: - informar a los estudiantes sobre sus				Respuestas abiertas sobre las técnicas de evaluación, agrupadas en: - tradicionales (pruebas y ejercicios)		Respuestas abiertas sobre la necesidad de cambio, agrupadas en: - cambiar para mejorar los

	<p>resultados y proceso</p> <ul style="list-style-type: none"> - para realizar informes (acreditar) - identificar estudiantes de bajo y alto rendimiento, y para dirigir el ritmo y el camino de su enseñanza - obligar a estudiar - para preparar para los exámenes 				<ul style="list-style-type: none"> - tradicionales y alternativas (investigación dirigida, juegos, portfolio...) 		<p>resultados de los alumnos</p> <ul style="list-style-type: none"> - no cambiar. Y sobre las preocupaciones, agrupadas en: - miedo al cambio (resistencias alumnos, aumento de la carga de trabajo...) - falta de imparcialidad y fiabilidad en las técnicas alternativas
<p>Blanco y González (2010) Educación Primaria.</p>	<p>Concepto de evaluación como actividad. Importancia, función y utilidad de la evaluación. Finalidades de la evaluación: inicial, continua y final.</p>	<p>Contenidos de la evaluación.</p>		<p>Criterios de evaluación. Temporalización de la evaluación: frecuencia de la evaluación inicial, continua y final. Modos de evaluación de las competencias básicas.</p>	<p>Instrumentos de evaluación para la evaluación inicial, continua y final. Elaboración de los instrumentos de evaluación.</p>	<p>Manejo del profesorado de los resultados de la evaluación: manera de darlos a conocer al alumnado e cómo comparte información sobre la evaluación a las familias.</p>	
<p>Alfageme et al. (2010) Secundaria</p>	<p>Finalidades o propósitos de la evaluación</p>	<p>Qué se evalúa.</p>		<p>Procedimientos de evaluación.</p>	<p>Instrumentos de evaluación.</p>	<p>Calificación.</p>	

Obligatoria. Ciencias Sociales.							
Rodríguez (2010) Educación Preescolar, Básica y Media. Comunicación lingüística	Finalidades / propósitos de la evaluación: - verificación y control - diagnóstico - prospectiva - auto-reflexión docente.		Tipo de evaluación según la participación de la comunidad: - heteroevaluación (centrada en el docente) - coevaluación (realizada entre pares) - autoevaluación (centrada en el estudiante) - con participación de los padres.		Uso de instrumentos: - rejillas - cuestionarios y guías - preguntas - no utiliza.	Uso de los resultados de evaluación: - constatación, corrección e implementación de refuerzos - retroalimentación y reflexión sobre aciertos y desaciertos - búsqueda de compromiso y responsabilidad del estudiante - rendición de cuentas sobre el rendimiento. - no se usan.	
Vergara (2011) Educación básica. Varias disciplinas.	Finalidad y/o funcionalidad de la evaluación: - en el sistema escolar - en el aula Sobre cuál es el concepto de evaluación del aprendizaje:	Objeto de la evaluación: - contenidos conceptuales - contenidos procedimentales - contenidos actitudinales-valóricos	Agente evaluador: - docente evaluador - alumno evaluador - evaluación de pares	Sistema de registro y análisis: - estrategias e instrumentos de evaluación - criterios de corrección y calificación - formas de evaluación		Sistema de calificación y promoción: - escala de calificación - reglamento de evaluación - reglamento de promoción	Evaluación y reforma: - relación evaluación-reforma - relación evaluación-aprendizaje

	<ul style="list-style-type: none"> - concepto de evaluación - actividad docente - proceso de evaluación 						<p>Factores que afectan el desempeño:</p> <ul style="list-style-type: none"> - relacionados al docente - relacionados al alumno
<p>Dolores y García-García (2017) Educación Secundaria. Matemáticas.</p>	<p>Respuestas abiertas sobre la finalidad, agrupadas en:</p> <ul style="list-style-type: none"> - medición de los conocimientos adquiridos - la demostración del desarrollo de competencias - saber qué problemas tiene el alumno con el aprendizaje <p>Sobre cómo definen la evaluación, agrupadas en:</p> <ul style="list-style-type: none"> - medición de los 	<p>Respuestas abiertas sobre qué se evalúa, agrupadas en:</p> <ul style="list-style-type: none"> - los temas o el plan de estudio - la valoración y medición del aprendizaje - el alumno - el compromiso de evaluar 	<p>Respuestas abiertas quiénes evalúa agrupadas en:</p> <ul style="list-style-type: none"> - solo el alumno - el profesor, los estudiantes y otros profesores - los padres de la familia y todo el personal de la institución 		<p>Respuestas abiertas sobre instrumentos, agrupadas en:</p> <ul style="list-style-type: none"> - exámenes - ejercicios, participación y tareas - otros: exposiciones, cuestionarios, asistencias y trabajos. <p>Y sobre actividades:</p> <ul style="list-style-type: none"> - tareas, ejercicios y participaciones - problemas y exposiciones - actividades ligadas a la realidad 		<p>Respuestas abiertas sobre qué es una actividad significaba, agrupadas en:</p> <ul style="list-style-type: none"> - aquellas con interés, motivación o atención para los alumnos, mayor significado o relevancia - aplicación del conocimiento a problemas reales - actividad de investigar

	conocimientos adquiridos. - los aprendizajes logrados - verificación o medición de competencias - criterios de evaluación						
--	--	--	--	--	--	--	--

Anexo 5. Cuestionario "Mi posicionamiento sobre evaluación..."

Mi posicionamiento sobre evaluación

El siguiente cuestionario es totalmente anónimo. Contiene mayoritariamente preguntas de respuesta única; en alguna de ellas puede costarte seleccionar solo una opción, pero te rogamos que elijas la que mejor te representa o te caracteriza. Muchas gracias por participar y contestar con sinceridad.

Dirección de correo electrónico *

Dirección de correo electrónico válida

Este formulario recopila las direcciones de correo electrónico. [Cambiar configuración](#)

¿Has participado en el proyecto GAPPISA? *

La evaluación auténtica: repensar la evaluación para enseñar competencias.

Sí

No

Datos sobre ti en GAPPISA

Descripción (opcional)

¿En qué curso/edición del Proyecto? *

*Por favor, indica el curso escolar (por ejemplo, 2013-2014)

Texto de respuesta corta

¿Finalizaste el proyecto? *

Sí

No

¿Eras coordinador/a? *

Sí

No

¿Sigues aplicando la prueba que diseñaste? *

- Sí, exactamente la misma
- Sí, pero la modifiqué para ajustarla al grupo
- No

¿Has diseñado nuevas pruebas, parecidas a esta, bajo alguna de las dimensiones de la guía GAPPISA? *

- Sí
- No

¿Has utilizado alguna prueba de un/a compañero/a del proyecto o de la base de datos GAPPISA? *

- Sí
- No

¿Quieres recibir asesoramiento sobre los nuevos materiales que has diseñado? *

*Nos pondremos en contacto contigo en la mayor brevedad posible.

- Sí
- No

Datos generales

Descripción (opcional)

Nivel educativo en el que impartes docencia *

1. Educación Infantil
2. Educación Primaria
3. Educación Secunda
4. Bachillerato
5. Ciclo formativo
6. Universidad
7. Otro

Escribir "Otro":

Docencia en (tutor/a, especialista de...): *

Texto de respuesta corta

Años de experiencia docente *

1. Menos de 5 años
2. De 5 a 20 años
3. Más de 20 años

En qué grado estoy satisfecho/a respecto a la evaluación que hago... *

1 2 3 4 5 6 7 8 9 10

Poca satisfacción Gran satisfacción

Cuestionario sobre evaluación educativa

Las siguientes preguntas deben responderse con respecto al momento actual.

1. A nivel de equipo docente y proyecto educativo... *

- a) se evalúa según las exigencias externas (pruebas de ciclo, de cambio de etapa...)
- b) existe un consenso previo sobre lo que significa evaluar y cómo se debe hacer.
- c) cada uno evalúa según sus concepciones y habilidades.
- d) se evalúa de acuerdo con un proyecto impuesto por una programación pre-diseñada.

2. Los instrumentos que utilizo con mayor frecuencia en mis actividades de evaluación son... *

- a) casos sobre problemas realistas que deben resolverse mediante la aplicación de competencias.
- b) ejercicios y problemas sobre conocimientos (conceptos y procedimientos) como los trabajados en clase.
- c) un registro de todas evidencias de aprendizaje, incluidas las actividades de evaluación dirigidas a analizar tanto ...
- d) pruebas tipo test o de modalidad de preguntas variadas de respuesta corta (relaciona, identifica, define, contesta...)

3. Lo que evalúo mayoritariamente son... *

- a) las competencias desarrolladas en los procesos de aprendizaje.
- b) los conocimientos (conceptos, procedimientos y actitudes) adquiridos por los alumnos.
- c) las estrategias utilizadas por los alumnos en las prácticas educativas.
- d) los aprendizajes de los alumnos y también mi actuación docente.

4. Cuando evalúo (desde el diseño de los instrumentos hasta la nota final), normalmente me siento... *

- a) con ganas de explicar y compartir lo que hago.
- b) presionado/a por si será bien aceptado.
- c) con ganas de probar su funcionamiento y mejorar.
- d) como un juez que debe dictaminar sobre quién realmente aprendió.

5. El momento en el que evalúo principalmente es... *

- a) a lo largo de la Unidad Didáctica o proyecto realizo actividades de evaluación para que los alumnos tomen consc...
- b) después de realizar todas las actividades programadas en la Unidad Didáctica o proyecto.
- c) antes de las actividades programadas en la Unidad Didáctica o proyecto.
- d) "in situ", durante las actividades de aprendizaje, para ajustar la ayuda educativa en cada momento.

6. Evalúo mayoritariamente con la finalidad de... *

- a) recoger indicadores que me ayuden a mejorar mi enseñanza.
- b) rendir cuentas del trabajo educativo realizado.
- c) acreditar unos niveles alcanzados por los alumnos.
- d) que mis alumnos se corrijan y mejoren sus aprendizajes.

7. Mis actividades y/o pruebas de evaluación cumplen las siguientes características de forma general *

Puedes marcar más de una. Considera que no es necesario que se cumplan todas las condiciones a la vez en un mismo instrumento y/o actividad de evaluación. Las opciones marcadas deben ser representativas de tu forma habitual de evaluar.

- a) los problemas que se presentan incluyen contenidos de varias materias.
- b) las tareas y/o preguntas son realistas, relevantes y socializadoras, y se refieren a una situación inicial a resolver ...
- c) existe una graduación en la exigencia cognitiva de las preguntas o actividades de evaluación.
- d) el formato de los enunciados y de las respuestas es heterogéneo (texto, imagen, audio, vídeos, blogs, actuación, ...
- e) las condiciones de resolución de la prueba son auténticas, es decir, similares a como se producirían en la realida...
- f) existen actividades optativas para que el alumno pueda escoger.
- g) hay preguntas sobre el proceso de resolución de la actividad y/o prueba por parte del alumno (planificación, regu...
- h) se proponen tareas de auto-evaluación y de co-evaluación.
- i) los criterios de evaluación y la forma de puntuar las actividades y/o pruebas están consensuados previamente co...
- j) la comunicación de los resultados y correcciones se entiende como una nueva oportunidad de aprendizaje y exist...

8. Una de las principales dificultades que tengo en relación a la evaluación es... *

- a) valorar las competencias adquiridas durante el proceso de aprendizaje.
- b) determinar una calificación final justa para cada alumno.
- c) que las actividades consigan que los alumnos sean conscientes de sus aprendizajes y sepan cómo y en qué mej...
- d) recoger indicadores para ajustar las ayudas educativas.

9. La persona que normalmente se encarga de evaluar es... *

- a) instancias externas (pruebas estandarizadas de la administración, pruebas planteadas en el libro de texto, etc.)
- b) el mismo alumno y/u otro alumno.
- c) yo como profesor exclusivamente en mis clases.
- d) múltiples agentes: yo, mis alumnos, otros profesores y personas externas al centro.

10. Escribe cómo mínimo uno y cómo máximo tres incidentes (situaciones que te han impactado emocionalmente) que te hayan pasado respecto a la evaluación.

Deben referirse a los últimos dos/tres años.

Incidente 1: *

Texto de respuesta larga

Incidente 2:

Texto de respuesta larga

Incidente 3:

Texto de respuesta larga

¡Muchas gracias por tu colaboración!

Anexo 6. Instrumento de análisis del posicionamiento docente sobre evaluación

DIMENSIÓN IDENTIDAD	VARIABLE DIMENSIÓN	EVALUACIÓN TRADICIONAL		EVALUACIÓN AUTÉNTICA competencial, formadora e inclusiva	
Concepciones	<i>Con qué finalidad se evalúa</i>	Acreditar los niveles alcanzados por los alumnos.	Rendir cuentas sobre la docencia impartida, que se traduce en los resultados de los alumnos	Recoger indicadores para mejorar la enseñanza y ajustar las ayudas educativas	Ofrecer ayudas educativas para que el alumnado pueda valorar sus aprendizajes y mejorar.
	<i>Qué se debe evaluar</i>	Evaluar conocimientos de forma equitativa y justa de acuerdo con una tabla de niveles preestablecidos.	Evaluar conocimientos , posiblemente relacionados con los objetivos competenciales expuestos en el currículum.	Realizar una auto evaluación del docente , a través del análisis de los resultados del alumnado y las estrategias de aprendizaje que utiliza.	Facilitar y acompañar una auto evaluación por parte del alumnado y la evaluación de las estrategias de aprendizaje que utiliza.
	<i>Cómo se debe evaluar</i>	Se debe partir de una programación diseñada por otros agentes para garantizar consenso entre los niveles alcanzados.	Se deben evaluar según lo exijan instituciones externas.	Cada docente debe evaluar su enseñanza para ajustar la ayuda según su criterio propio.	El equipo docente debe consensuar y establecer unas líneas comunes para evaluar las estrategias de aprendizaje del alumnado y que éste se evalúe.
Estrategias	<i>Quién de evaluar</i>	El responsable final de la evaluación es una institución o equipo que establece cuáles han de ser los niveles preestablecidos.	El responsable final de la evaluación es la escuela que responde a las instituciones externas y los criterios para rendir cuentas.	El principal responsable de evaluar es el docente puesto que él es el que mejor conoce el alumnado.	El principal responsable de evaluar son los propios alumnos mediante el uso de los instrumentos que aporta el docente.
	<i>En qué momento se debe evaluar</i>	Se debe evaluar al final de la Unidad Didáctica o Proyecto.	Se debe evaluar al final de la Unidad Didáctica o Proyecto.	Se debe evaluar durante y después de la Unidad Didáctica o Proyecto y, especialmente, “in situ” , durante las actividades para poder ajustar la ayuda educativa.	Se debe evaluar a lo largo (antes, durante y después) de la Unidad Didáctica o Proyecto y, especialmente, antes para que los alumnos puedan valorar su aprendizaje a través de la comparación de evidencias.
	<i>Qué instrumentos se deben usar</i>	Se deben aplicar exámenes de respuesta única o de diferente tipología que miden los conocimientos.	Se deben aplicar exámenes basados en ejercicios y problemas que miden los conocimientos que se	Se deben utilizar instrumentos de evaluación basados en la resolución activa de problemas auténticos donde	Se deben utilizar instrumentos relacionados con el análisis y reflexión sobre el aprendizaje a partir de las evidencias creadas

			relacionan con los objetivos competenciales.	el docente pueda valorar las estrategias de aprendizaje y la eficacia de las ayudas educativas.	durante la resolución activa de problemas auténticos (carpeta de aprendizaje, rúbricas, bases de orientaciones, etc.).
Sentimientos	<i>Con qué motivación se debe evaluar</i>	La acción está asociada a una motivación relacionada con impartir justicia .	La acción está asociada a una motivación relacionada con demostrar el trabajo realizado .	La acción está asociada a una motivación relacionada con la mejora y constante actualización de la actividad docente.	La acción está asociada a una motivación relacionada con las ganas de compartir estrategias de aprendizaje para que el alumnado pueda evaluarse.
	<i>Qué dificultades deben superarse</i>	La problemática principal está relacionada con adjudicar una calificación justa .	La problemática principal está relacionada con saber valorar las competencias adquiridas como indicador de resultados de centro .	La problemática principal está relacionada con recoger información precisa para ajustar la ayuda educativa , considerando la cantidad de alumnos que hay.	La problemática principal está relacionada con aportar evidencias e instrumentos útiles para que el alumnado pueda aprender con la evaluación.

