



ISSN: 1989-0397

**Revista
Iberoamericana
de Evaluación Educativa**

**Mayo 2016 - Volumen 9, número 1
doi: 10.15366/rie2016.9.1**

**rinace.net/rie/
revistas.uam/rie**



CONSEJO EDITORIAL

DIRECTOR

F. Javier Murillo

EDITORA

Nina Hidalgo Farran

ASESOR EDITORIAL

Manuel Lorite Becerra

CONSEJO DIRECTIVO

Marcela Gajardo (Programa de Promoción de la Reforma Educativas de América Latina y El Caribe, PREAL)

Sergio Martinic (Pontificia Universidad Católica de Chile)

Carlos Pardo (Instituto Colombiano para la Evaluación de la Educación, ICFES)

Margarita Poggi (Instituto Internacional de Planeamiento de la Educación -IIPE-. UNESCO, Argentina)

Francisco Soares (Universidade Federal de Minas Gerais, Brasil)

CONSEJO CIENTÍFICO

Juan Manuel Álvarez. Universidad Complutense de Madrid, España.

Patricia Arregui. Grupo de Análisis para el Desarrollo (GRADE), Perú.

Daniel Bogoya. Universidad Pedagógica Nacional, Colombia.

Nigel Brooke. Universidade Federal de Minas Gerais, Brasil.

Leonor Cariola. Ministerio de Educación, Chile.

María do Carmo Clímaco. Universidade Lusófona de Humanidades e Tecnologias (ULHT), Portugal.

Cristian Cox. Pontificia Universidad Católica de Chile.

Santiago Cueto. Grupo de Análisis para el Desarrollo (GRADE).

Tabaré Fernández. Universidad de la República, Uruguay.

Juan Enrique Froemel. Universidad UNIACC, Chile.

Reyes Hernández-Castilla. Universidad Autónoma de Madrid, España.

Rubén Klein. Fundação Cesgranrio, Brasil.

Luis Lizasoain. Universidad del País Vasco/Euskal Herriko Unibertsitatea, España.

Jorge Manzi. MIDE-UC, Pontificia Universidad Católica de Chile.

Joan Mateo. Universidad de Barcelona, España.

Liliana Miranda. Ministerio de Educación de Perú.

Carlos Muñoz Izquierdo. Universidad Iberoamericana Ciudad de México, México.

Margarita Peña. Instituto Colombiano para la Evaluación de la Educación, ICFES.

Dagmar Raczynski. Asesorías para el Desarrollo, Chile.

Héctor Rizo. Universidad Autónoma de Occidente, Colombia.

Mario Rueda. Universidad Nacional Autónoma de México.

Guadalupe Ruíz. Universidad Autónoma de Aguascalientes, México.

Ernesto Schiefelbein. Universidad Autónoma de Santiago, Chile.

Alejandra Schullmeyer. Instituto Nacional de Estudios Pedagógicos, Brasil.

Javier Tejedor. Universidad de Salamanca, España.

Flavia Terigi. Universidad de Buenos Aires, Argentina.

Alexander Ventura. Universidade de Aveiro, Portugal.

ÍNDICE

Editorial

XXX	5
<i>F. Javier Murillo y Nina Hidalgo</i>	

Temática libre

Conceptions and Practices of Assessment in Higher Education: A Study of Portuguese University Teachers	9
---	---

Diana Ribeiro Pereira y Maria Assunção Flores

La Evaluación a través de Portafolios: ¿Una Ocasión para el Aprendizaje?	31
---	----

Noelia Alcaraz Salarirche

Análisis de una Escala para Evaluar la Satisfacción del Profesional Docente en la Enseñanza Primaria	47
---	----

Rosario Josefa Rodríguez Pulido, Jaime León González-Vélez, María Victoria Aguiar Perera y Josué Artilés Rodríguez

Evaluación Psicopedagógica e Intervención Sociocomunicativa en Niños con Trastorno de Asperger. Análisis de un Caso	61
--	----

Cecilia Latorre Cosculluela y Miguel Puyuelo Sanclemente

Una Noción de Evaluación: La Evaluación Formativa en el Marco de las Concepciones de Finlandia y la Ciudad de Buenos Aires	75
---	----

Celina Cardoner Zunino

Concepciones sobre Evaluación en la Formación Inicial del Profesorado de Ciencias, Tecnología y Matemáticas	91
--	----

Alfonso Pontes Pedrajas, Francisco J. Poyato López y José M^a. Oliva Martínez

El Papel de la Evaluación a Gran Escala como Política de Rendición de Cuentas en el Sistema Educativo Mexicano	109
---	-----

José Alfonso Jiménez Moreno

Problemas Éticos y Representaciones Sociales de la Evaluación en la Universidad	127
--	-----

Nelly Patricia Bautista Cárdenas

Funcionamiento Diferencial del Ítem en una Evaluación Estandarizada según Necesidades Educativas Especiales Transitorias 145

Fabiola Silva Morales y María Verónica Santelices Etchegaray

El Programa de Evaluación Docente en la UADY. Experiencia de la implementación 161

Norma Alicia Benois Muñoz, Fabiola Isabel Gómez Cauich, Maritza Minelli Briceño Caballero y Jessica Betzabé Zumárraga

Editorial

Evaluación Democrática y para la Democracia

F. Javier Murillo * y Nina Hidalgo **

Universidad Autónoma de Madrid

En un libro publicado unos días antes de morir, el sociólogo francés Pierre Bourdieu (2001) destacó el carácter claramente anti-democrático de la evaluación de los estudiantes. En línea con filósofos críticos como Michel Foucault (2001), Bourdieu consideró que la evaluación ejerce un poder coercitivo simbólico hacia los alumnos y alumnas a través de mecanismos tales como el uso de pruebas externas estandarizadas que sirven más para ordenar y clasificar que para mejorar, la opinión pública que se genera a partir de los resultados de evaluación cuando se traducen en el uso de rankings, la dependencia hacia la evaluación que se crea entre los estudiantes ya desde edades tempranas, o el falso mito de que las calificaciones reflejan el aprendizaje de los estudiantes.

Aunque en estos 15 años transcurridos se han multiplicados los esfuerzos, especialmente teóricos, para ofrecer alternativas en la evaluación, como el surgimiento de enfoques tales como la Evaluación Auténtica, la Evaluación Culturalmente Sensible, la Evaluación Crítica, la Evaluación Democrática Deliberativa, o la Evaluación para la Justicia Social (Murillo e Hidalgo, 2015), la realidad muestra que sigue ejerciendo un fuerte poder coercitivo sobre los estudiantes y sobre los docentes, convirtiéndose en el principal instrumento de control educativo, especialmente por parte de las Administraciones (Shohamy, 2001). Así, no hay forma más eficaz de controlar a los profesores y profesoras, a los estudiantes y a lo que acontece en las aulas y escuelas -poniendo incluso en duda la capacidad y el profesionalismo de los trabajadores de la educación-, que imponer sistemas externos de evaluación con consecuencias vitales para el futuro de las escuelas y de los propios niños, niñas y adolescentes.

Sin que lo anterior deje de ser cierto, no lo es menos que la evaluación es una de las herramientas más potentes de las que disponemos para romper con prácticas educativas autoritarias y convertirse en una herramienta democrática para el aprendizaje de los estudiantes y para el cambio social. Efectivamente, su situación privilegiada dentro del proceso de enseñanza y aprendizaje, pero también ser un elemento crítico que mediatiza y es mediatizado por la cultura escolar, le hace que pueda convertirse en un potente catalizador que genere un proceso de transformación de las prácticas docentes y la cultura escolar. Una evaluación democrática puede contribuir a la democratización de las aulas, de la escuela y, por qué no, de la sociedad.

Seamos claros. ¿Es posible una sociedad democrática con escuelas autoritarias? ¿Es posible tener escuelas y aulas democráticas si las evaluaciones son represoras, jerárquicas y coercitivas en su diseño, desarrollo, corrección y devolución? Nos cuesta imaginarlo.

Sin pretender elaborar un manual, sí que nos gustaría aportar algunas ideas para ayudarnos a reflexionar sobre el sentido de la Evaluación Democrática de los estudiantes:

- ✓ *Empoderar a los estudiantes.* La evaluación debe convertirse en un proceso que empodere a los alumnos y alumnas en su propio proceso de aprendizaje. Una práctica evaluativa democrática requiere la participación activa de los estudiantes en todo el proceso evaluativo: en el diseño, en la ejecución, en la corrección, en la devolución de la información y en las decisiones que se toman derivadas de la evaluación (Greene, 2000). No supone únicamente conocer sus opiniones o tomar alguna decisión a través de una votación, sino convertirles en responsables últimos de su aprendizaje.
- ✓ *De la evaluación de los estudiantes a la evaluación con ellos.* Nadie cambia si no quiere hacerlo. Estrategias de autoevaluación, que busquen una reflexión personal del estudiante de su propio aprendizaje así como de los retos a superar, es una estrategia democrática y eficaz para avanzar.
- ✓ *Crítica.* La evaluación también ha de posibilitar a los estudiantes reflexionar críticamente y aportar sus puntos de vista alternativos, creativos y críticos. La búsqueda de la “verdad única” que el docente posee y el alumno desconoce, y que solo puede alcanzarse por el método prescrito es como suena, una práctica represiva.
- ✓ *Justa.* Una evaluación nunca podrá ser democrática sino es justa. La justicia supone una atención diferencial a los estudiantes, dando más a aquellos que más lo necesitan y utilizando estrategias variadas para no favorecer a un determinado tipo de alumno o alumna. De igual forma, ha de contribuir al desarrollo integral de los estudiantes, evaluando no únicamente los aspectos cognitivos sino también actitudinales, procedimentales, afectivos... Pero de nada sirve todo esto si no valoramos el avance real de los estudiantes. El recorrido que cada estudiante realiza es individual y por lo tanto, la evaluación también tiene que serlo.
- ✓ *No jerárquica.* La práctica evaluativa ha de estar enmarcada en una reconsideración de la figura y el papel del docente que se libera de su armadura de poder para ser un mediador en el proceso de enseñanza y aprendizaje. Siguiendo a Freire (1971), hay que superar la idea de que el que “sabe” deposita en el que “no sabe” el conocimiento, ya que esto supone un mantenimiento de las estructuras de poder existentes tanto en la escuela como en la sociedad.
- ✓ *Cooperativa.* El trabajo en equipo y la cooperación entre los estudiantes ha de formar parte de la esencia de la evaluación democrática. Enseñar a los estudiantes a trabajar de forma conjunta, colaborativa y participativa favorece el desarrollo de una evaluación democrática. Este trabajo cooperativo en términos evaluativos, no debe realizarse únicamente entre los estudiantes, sino también entre los docentes y con familias. La comunidad es un elemento esencial de la democracia, y debe estar presente y activa en las escuelas.

- ✓ *Social*. De acuerdo con lo que afirman House y Howe (2000, p. 3) “las prácticas evaluativas están firmemente arraigadas a determinadas estructuras sociales e institucionales”, y por lo tanto tienen que tener el poder de transformar la sociedad. Una vez la evaluación es democrática, es decir, se desarrolla a través de procesos democráticos donde participan a los estudiantes, es necesario que la misma prepare a los estudiantes para favorecer la democracia en la sociedad. Así, dotarles de herramientas para hablar en público, ser críticos con la realidad, expresar sus ideas y conocer los cauces sociales de participación es fundamental para que la evaluación suponga un cambio, no tan solo en el aula, sino especialmente en la sociedad.

En las escuelas, los estudiantes aprenden cómo funciona la sociedad así como su papel en la misma. A través de la evaluación, los docentes pueden concienciar a los estudiantes acerca de su papel crítico en la transformación social, ayudarles a comprender que su participación social es fundamental igual que lo es en su aprendizaje. Si la escuela quiere cambiar la sociedad es necesario que empiece por incluir a los estudiantes en su evaluación, enseñándoles a tomar decisiones y ser responsables de su aprendizaje.

Transformar las prácticas evaluativas para que promuevan la democracia en la sociedad supone dejar de ver la evaluación como un simple instrumento de medición y comprenderlo como una herramienta al servicio del aprendizaje autónomo de los estudiantes y de la conformación de una sociedad más justa y democrática.

Referencias

- Bourdieu, P. (2001). *Langage et pouvoir symbolique*. París: Seuil/Points.
- Foucault, M. (2001). *Vigilar y castigar*. Madrid: Siglo XXI Editores.
- Freire, P. (1971). *Pedagogía del oprimido*. Ciudad de México: Siglo XXI Editores.
- Greene, J. C. (2000). Challenges in practicing deliberative democratic evaluation. *New Directions for Evaluation*, 2000(85), 13-26.
- House, E. R. y Howe, K. R. (2000). Deliberative democratic evaluation in practice. En D. Stufflebeam, G. Madaus y T. Kellaghan (Eds.), *Evaluation models* (pp. 409-421). Londres: Springer.
- Murillo, F.J. e Hidalgo, N. (2015). Enfoques Fundamentantes de la Evaluación de Estudiantes para la Justicia Social. *Revista Iberoamericana de Evaluación Educativa*, 8(1), 43-61.
- Shohamy, E. (2001). Democratic assessment as an alternative. *Language Testing*, 18(4), 373-391.

Conceptions and Practices of Assessment in Higher Education: A Study of Portuguese University Teachers

Concepciones y Prácticas de Evaluación en la Educación Superior: Un Estudio con Profesores Universitarios Portugueses

Diana Ribeiro Pereira* y Maria Assunção Flores
Universidade do Minho

This paper looks at how Portuguese university teachers look at assessment in Higher Education. It focuses on their conceptions and methods of assessment. Data were collected through face to face interviews and online open-ended questionnaires in five Portuguese Public Universities in different fields of knowledge. In total, 57 teachers participated in this study. Findings are presented according to the emerging categories arising from the data analysis: i) university teachers' conceptions of assessment; ii) most used assessment methods; iii) role of assessment; iv) key moments in which assessment is put into practice. Implications of the findings are discussed.

Keywords: Assessment, Higher Education, Teaching, Learning, Assessment methods, Conceptions of assessment.

Este artículo analiza cómo los profesores universitarios portugueses ven la evaluación en la Educación Superior. Se centra en sus concepciones y métodos de evaluación. Los datos fueron recolectados mediante entrevistas cara a cara y cuestionarios *online* de composición abierta en cinco Universidades Públicas Portuguesas en diferentes campos del conocimiento. En total, 57 profesores participaron en este estudio. Los resultados se presentan según las categorías emergentes que surgen del análisis de los datos: i) las concepciones de la evaluación de los profesores universitarios; ii) el papel de la evaluación; iii) los momentos clave en que se utiliza la evaluación; iv) los métodos de evaluación más usados. Se discuten las implicaciones de los datos.

Palabras clave: Evaluación, Educación Superior, Enseñanza, Aprendizaje, Métodos de evaluación, Concepciones de evaluación.

1. Introduction¹

Assessment in higher education impacts on students' learning in various ways (Ramsden, 2004; Gibbs, 1999; Scouller, 1998; Simms & George, 2014). The ways in which assessment is carried out have implications for students' learning (Segers, Gijbels, & Thurlings, 2008) and impact on them even before they are assessed (Rowntree, 1987), namely in terms of their approaches to learning (Boud & Falchikov, 2006; Fletcher, Meyer, Anderson, Johnston, & Rees, 2012). The ways in which teachers see assessment influence their assessment practices and how they relate these practices to the teaching and learning process (Fletcher et al., 2012; Samuelowicz & Bain, 2002). Teachers' approaches to teaching are also related to their conceptions of teaching (Trigwell & Prosser, 1996) which will influence in turn how they assess students' learning (Watkins, Dahlin, & Ekholm, 2005). Earlier studies show how different conceptions of teaching and learning impact on students' assessment. Prosser and Trigwell (1998) identify two conceptions of teaching and learning: 1) teachers who look at learning as joining of information view teaching as transmission of information to students and are more focused on themselves as teachers; and 2) teachers who see learning as transforming the students' conceptions view teaching as helping, developing and changing these conceptions and thus are more student-focused. Samuelowicz and Bain (2002, p. 181) found that "different teachers' orientations result in different assessment practices". Teachers who see the teaching and learning process as reproduction or transmission of knowledge view assessment as students' ability to reproduce the knowledge acquired. In contrast, teachers who see teaching as facilitating learning and promoting critical thinking view assessment as transformation of knowledge and as an integral part of the learning process (Samuelowicz & Bain, 2002).

This assumption is in line with existing literature. Ramsden (2004), for instance, reports that teachers who see teaching as knowledge transmission see assessment as a separate element of teaching instead of an integrated approach that promotes deep approaches to learning. Postareff and Lindblom-Ylänne (2008) found that teachers who see teaching as construction of knowledge tend to adopt several and different assessment practices rather than just the traditional ones. This may be related to another study of Postareff, Virtanen, Katajavuori and Lindblom-Ylänne (2012) which indicates that the majority of teachers use summative assessment and traditional methods and a minority uses formative assessment and a range of different methods.

These different perspectives show how conceptions of teaching and learning may influence teachers' attitudes towards assessment. In fact, teachers' conceptions of assessment may influence their teaching practices (Brown, 2004; Brown, Lake, & Matters, 2011; Pajares, 1992) and students' learning (Brown & Hirschfeld, 2008; Marton & Saljo, 1997).

Remesal (2011) proposed a model about the conceptions of assessment regarding four different dimensions: 1) learning process; 2) teaching process; 3) accreditation of

¹ This work was supported by the Portuguese Foundation for Science and Technology under the Grant SFRH/BD/76175/2011.

learning and 4) accountability of the professional teaching activity. In this paper, the term conceptions of assessment is used in relation to teachers' views about assessment concerning the dimension of the learning process, the teaching process and the accreditation of learning since they are aligned with the focus of this study.

Although there are studies focusing on teachers' and students' conceptions of assessment and their impact on learning (Brown & Hirschfeld, 2008; Fletcher et al., 2012; Gibbs & Simpson, 2004) more needs to be known about teachers' beliefs related to assessment and their assessment practices, as well as the relationship between their beliefs and the processes of teaching, learning and assessment (Samuelowicz & Bain, 2002). Accordingly, it is important to get to know the ways in which university teachers' look at assessment, because their conceptions of teaching, learning and assessment influence how they teach and how students learn (Brown, 2004; Thompson, 1992).

This study intends to illustrate the conceptions of Portuguese university teachers on the assessment process after the so-called Bologna process. The reorganisation of higher education systems in Europe has challenged the ways of looking at curricula but also to teaching, learning and assessment (Flores & Veiga Simão, 2007; Flores, Veiga Simão, Barros, & Pereira, 2015; Simão, Santos, & Costa, 2003). Among the key features of the Bologna Process is the emphasis on a student-centred approach involving new forms of teaching and learning with tutorial support (Leuven / Louvain-la-Neuve Communiqué, 2009) and the use of innovative methods that promote participatory learning and the development of critical thinking (Bucharest Communiqué, 2012). However, these changes may not occur in all higher education systems in the same way. Therefore, as the context of this study is part of the European Area of Higher Education (EAHE), implemented by the Bologna Process, this study may contribute to get to know the university teachers' perceptions of assessment as to compare and contrast them with findings obtained in other contexts.

2. Conceptions of Assessment

Assessment has been investigated in all levels of teaching. It has developed according to different understandings and conceptions such as *assessment as a tool for learning* (Dochy & McDowell, 1997) *as improvement for teaching and learning* (Brown et al., 2011) and *learner-centred assessment* (Webber, 2012) moving beyond the perspective of accreditation and certification. These conceptions focused on the potential benefits that assessment can bring to the learning process (Dochy & McDowell, 1997) in its formative nature (Brown et al., 2011) and in the use of assessment methods that look at student as learner (Webber, 2012). The assessment functions may also determine the learning process. Hadji (1994) suggests that assessment has three main functions, corresponding to three main different goals: the purpose of *certification* associated with *summative assessment*; the purpose of *regulation* associated with *formative assessment* and the purpose of *guidance* associated with *diagnostic or prognostic assessment*. The certification, often seen as the traditional function of the assessment, assures the potential employer that a given student obtained the required training by recognising his/her competencies. However, this provides vague details of the kinds of knowledge and skills acquired by the student and the level of mastery in each one. It only informs about what the student knows in general and if the student reached the goals to pass or to start a profession (Perrenoud, 1999). In addition, this kind of assessment develops

hierarchies of excellence, establishing the progression of a cycle of studies and students are usually compared among themselves (Boud, 2000; Fletcher et al., 2012; Hadji, 1994) by obtaining a grade (Perrenoud, 1999). Thus, summative assessment entails a sum of a course, performed at a given time, being an assessment that measures results achieved by the students (Light & Cox, 2003). Some authors suggest that summative assessment is opposed to formative assessment (Bloom, Hastings, & Madaus, 1971; Light & Cox, 2003) since the goal is to make a very general assessment of the degree in which objectives have been achieved during the course or for any substantial part of it, usually taking place at the end of a period of the year to assign a note and subsequently a certificate (Brew, Riley, & Walta, 2009). In contrast, formative assessment has a pedagogical propose and it is intrinsic to the process of teaching itself (Brown & Knight, 1994; Hadji, 1994). The formative role of assessment is important to improve learning, because it provides students with feedback during the process of learning (Brown, Bull, & Pendlebury, 1997). It also enables them to have opportunities to improve (Brown, Race, & Rust, 2004) making it possible for both students and teachers to get to know how learning is proceeding (Biggs, 2003) essentially its successes and difficulties.

Biggs (2003) makes a clear distinction between the general purposes of these two types of assessment: formative and summative. The author suggests that summative assessment is labelled often as a negative assessment, since it creates surface approaches to learning and the students use less their cognitive skills (less than currently required), which results in a fragmented learning. In regard to formative assessment, the feedback that results from it facilitates learning, promotes the necessary information for deep approaches to learning resulting in a further development of cognitive activities that are appropriate to the level of the proposed task. However, Black and William (1998) found that formative assessment also entails some weaknesses when it is put into practice in contexts in which there is lack of awareness among teachers about the formative purpose of assessment. The authors also found that teachers value most the grading function than the learning function (Black & William, 1998).

3. Assessment Methods in Higher Education

The assessment methods adopted by university teachers have an important role in the quality of learning (Atkins, 2004; Fernandes, Flores, & Lima, 2012; Flores et al., 2015; Hue, Leung, & Kennedy, 2014; MacLellan, 2004; Pereira, Flores, & Niklasson, 2015). Several factors are influential of the most used assessment methods, either negatively or positively, especially on student learning and may be seen as an incentive for study and for improved performance (Biggs, 2003; Boud & Falchikov, 2007; Brown & Knight, 1994; Brown et al., 1997; Watering, Gijbels, & Dochy, 2008). The ways in which students look at learning are influenced by the ways in which they perceive assessment tasks (Brown & Knight, 1994; Drew, 2001).

Also, teaching methods must be aligned with assessment methods, taking into account the learning goals in order for teaching to be more effective (Biggs, 2003). While teachers see the objectives of the curriculum as key elements in the teaching and learning process, students look mainly at the ways in which assessment is carried out (Biggs, 2003; Ramsden, 2004). For this reason assessment cannot be seen as the end of the process (Dochy & McDowell, 1997) in so far as students pay attention at it at first

and then based on that they create a defined representation of the curriculum and activities in which they are to be involved (Meyers & Nulty, 2009). The use of given assessment methods and their adequacy to teaching and learning goals are thus of paramount importance. However, traditional methods frequently used in higher education, such as the exam or written test, while they are effective in some contexts and for given purposes, may not be suitable for all assessment purposes and may encourage reproduction and memorisation (Biggs, 2003; Pereira & Flores, 2012; Perrenoud, 1999). In fact, existing research shows that written tests promote low levels of comprehension (Dochy, Segers, Gijbels, & Struyven, 2007), reproduction of information under pressure and surface approaches to learning (Brown et al., 1997). On the other hand, the so-called alternative assessment methods or learner-centred methods (Webber, 2012), such as portfolios, projects, self- and peer assessment, simulations, collaborative assessment, among others (Struyven et al., 2005; Flores et al., 2015; Pereira et al., 2015) promote collaborative learning, and together with self and peer assessment seem to be more effective regarding deep learning and the development of new skills and professional attitudes. These methods also enable a more effective learning (Birenbaum & Feldman, 1998; Sambell & McDowell, 1998; Struyven, Dochy, & Janssens, 2005; Tang, Lai, Arthur, & Leung, 1999), fostering the development of autonomy, sense of responsibility, and reflection (Sambell & McDowell, 1998) and influencing the ways in which students see their own learning in a more positive way (Sluijsmans, Dochy, & Moerkerke, 1998). These methods also provide students with feedback about their performance (Brown et al., 1997) and prepare them to workplace situations (Biggs, 2003). Other studies, however, suggest that the non-traditional methods do not always change the perceptions of students nor lead to deep learning (Segers et al., 2008). As such, the different approaches to learning may be influenced by the assessment methods and assessment tasks used (Struyven et al., 2005; Fletcher et al., 2012) but also by the contexts in which they are used. Often, the problem of students having a surface approach to learning has to do with assessment tasks which relate to teachers' practices and the alignment with the aims of teaching and its environment (Biggs, 2003). Thus it is important to investigate the conceptions of assessment and the practices used by university teachers in different fields of knowledge.

4. Methods

This paper draws upon a wider piece of research within the context of a PhD in Educational Sciences, with a grant from the Portuguese Foundation for Science and Technology (SFRH/BD/76175/2011). The following questions are addressed in this paper:

1. How do university teachers look at assessment in Higher Education (HE)?
2. What kinds of assessment methods are used in HE? And why?
3. How do university teachers relate assessment methods and issues of teaching and learning?

4.1. Participants

The participants in this study are university teachers teaching year 3 in five Portuguese Public Universities. In total 57 teachers participated in the study (see table 1); 53% are male and 47% are female teachers. Their age ranged between 30 and 68

years old. Most of them (58%) are Assistant Professors and the years of teaching experience in HE ranged between 2 and 44 years. Most of them have 29 years of teaching. Out of 57 teachers 32 teach in Social Sciences and Humanities, 10 in Life and Health Sciences, 9 in Natural and Environmental Sciences and 6 in Sciences and Engineering.

Out of the 57 teachers 24 were face-to-face interviewed and 33 teachers responded to the questions using the link provided via email. In this study the four scientific fields of research identified at the Portuguese Foundation for Science and Technology were used: Social Sciences and Humanities (SSH), Life and Health Sciences (LHS), Natural and Environmental Sciences (NES) and Sciences and Engineering (ESE). Different programmes were selected: SSH: (Educational Sciences, Basic Education and Economics); LHS: (Nursing, Medicine, Pharmacy); NES: (Biology, Geology) and SE: (Mechanical Engineering, Computers Engineering and Biochemistry) which were operating in the five public universities.

Table 1. Participants

FIELD OF KNOWLEDGE		F
Social Sciences and Humanities	Educational Sciences	17
	Economics	12
	Basic Education	3
	Total	32
Life and Health Sciences	Nursing	5
	Medicine	4
	Pharmacy	1
	Total	10
Natural and Environmental Sciences	Biology	9
	Total	9
Sciences and Engineering	Mechanical Engineering	4
	Computer Engineering	1
	Biochemistry	1
	Total	6
Total		57

Source: Prepared by the authors.

4.2. Data collection and analysis

Data were collected between October 2012 and June 2013 in five Portuguese Public Universities. A research protocol was sent to the Presidents of Faculties and Institutes and to the Presidents of the Pedagogical Council of each Faculty and Institute in order to request permission to conduct the study. After obtaining the authorisation to conduct the study directors of the different programmes were contacted in order to obtain the email contacts of the teachers who were teaching in the third year of each programme. The interview protocol was validated with teachers teaching in one Public university between March and May 2012.

Face to face interviews and responses via email were received from 10 out of the 14 programmes selected. The participants were identified within the different departments in each programme. They were invited to participate in a face-to-face interview. Some of them were very busy and preferred to participate by sending their responses via email. For that a link with open-ended questions were created which was sent to them via email. Informed consent and confidentiality were granted for both the interviews and responses sent via email. The interview protocol was designed to obtain

data on the conceptions of assessment, assessment methods used in HE and connections between assessment, teaching and learning (see appendix 1).

This study is based on the interpretative paradigm which focuses on the interpretation of a phenomenon by understanding the meanings through the experiences of the individuals in a constantly changing reality (Blaikie, 2010; Cohen, Manion, & Morrison, 2008). Within this framework, through the use of qualitative methods the researcher has a systematic and integrated idea of the context and the holistic dimension in order to obtain a complex and substantial information (Miles, Huberman, & Saldaña, 2014). Therefore, a qualitative research design was adopted as it allowed exploring university teachers' views enabling a deeper understanding of the phenomenon (Silverman, 2013) and its complexity (Miles & Huberman, 1994). The interview was chosen because it enabled a broad retrospective, a projection of the actions, feelings, experiences of individuals and knowledge (Kvale, 1996). Interviews were transcribed verbatim. Content analysis was used to identify emerging categories, enabling a systematic description, through the categorisation of data (Schreier, 2012). Content analysis as a technique to analyse data performs valid inferences, leading to the emergence of what is relevant (Weber, 1990) and answer to research questions (Thomas, 2003).

5. Findings

Findings are presented according to the emerging categories arising from the data analysis: i) university teachers' conceptions of assessment; ii) most used assessment methods; iii) role of assessment; iv) key moments in which assessment is used.

5.1. University teachers' conceptions of assessment

The following section describes the findings regarding the conceptions of assessment in terms of *better assessment for students to learn*. Teachers spoke of continuous assessment during the process, formative assessment, and assessment based on active and participatory methods as the kinds of assessment that may help students learn better. Other participants do not point to any kind of assessment method as being better; some of them state that university teachers should use different methods to assess depending on the course or module.

Most of the participants claim that continuous assessment helps the students to learn better "together with independent work of the student with individual feedback" (P.4 SSH).

Some of them argue that this kind of assessment brings benefits to learning, namely if it is based "on a continuous logic that promotes a better quality of learning" (P.24 SSH). Also, continuous assessment is said to "promote critical thinking" (P.49 LHS).

Some university teachers believe that combining continuous assessment with formative assessment is positive for students' learning: "A continuous assessment, more personalised and adapted to each case, more formative and with possibility of reconstruction of knowledge is better" (P.25 SSH).

However, some of them point to difficulties in putting continuous assessment in practice in universities due to the number of students and heavy workload: "The ideal is a continuous assessment but it is impossible (P.37 NES) and "Continuous assessment

is the ideal, but there is a problem, it is hard to do continuous assessment in a class with 90 students” (P.45 LHS).

Formative assessment is seen, by some teachers, as important to students’ learning: “The formative assessment is the best assessment form. With formative assessment the students have to face situations, have to share and are always being monitored” (P.56 SSH).

However, like continuous assessment, formative assessment is difficult to be put into practice:

I have no doubts that formative assessment is the best assessment for students’ learning. The entire reproductive model evokes in the student an intellectual detachment, because the students do not participate and are not motivated, so the traditional model is not good... the formative assessment model needs more teachers, more resources, more technology and our country does not have fund for it. (P.47 SSH)

Assessment based on practical and active methods are seen as important for learning and improvement: “Assessment has to be combined with goals and methodologies of work, and the more effective is the use of active methods (P.22 SSH).

The importance of the practical context in assessment is also highlighted.

It is important to promote long-term learning and reasoning. Students will be future doctors and will be faced with questions on a daily basis, they will have to think and reflect, and if we use an assessment methodology that only point to memorisation they will not be prepared. The best assessment method will be the one that is based on practical things and prepare them for the practical context. (P.46 LHS)

Teachers emphasise continuous assessment and formative assessment as the best assessment to help students to learn better because it allows feedback and the reconstruction and regulation of learning. Assessment based on active and participatory methods is also pointed out by the participants. However, a number of constraints hinder the use of these assessments due to the number of students or to the number of teachers or to reasons related to available resources as can be seen in the following teacher account:

I have introduced before new active methods such as PBL. However, I think that Bologna is not an effective process, because not everybody shares the principles of Bologna, not everyone discusses these changes. I do not see how Bologna can work in classes with 100 students (...) governments that signed this document are not consistent (...) we face reductions in all resources. Education should be individualised because each student is different, and in our public universities this does not exist, we are talking about mass education. (P.43 CVS)

It is also claimed that there is a lack of conditions to put into practice some of the demands of the Bologna Process, as a student centred-approach, continuous assessment and the skills’ development through assessment practices. Furthermore, the Bologna Process “has followed different paces, across and within countries, and that the coexistence of the old and new systems in some countries has most likely favoured the emergence of controversy regarding its implementation in higher education institutions, students and labour markets” (Portela, Sá, Alexandre, & Cardoso, 2009, p. 466). These factors may explain, at least in part, some resistance to follow the assumptions of the Bologna Process, particularly with regard to the teaching, learning and assessment process, appearing not exist a common and shared view of the teaching and assessment practices across the European countries.

5.2. Most used assessment methods

Findings indicate that 41 out of the 57 teachers interviewed used written tests (see table 2). Practical work in groups, oral presentations, individual work and reports are also identified by some university teachers. Other assessment methods, although less used, included written reflections, interviews, literature review, reading tasks, etc.

Table 2. Assessment methods most used by the participants

	N	SSH (N=32)	LHS (N=10)	NES (N=9)	ESE (N=6)
Tests/Exam	41	18	10	8	5
Practical work in group	18	12	2	3	1
Oral Presentations	15	6	2	4	3
Individual work	10	6	2	2	0
Report	8	3	2	1	2

Source: Prepared by the authors.

As observed in table 2, written test or exam is one of the most recurring methods. These finding corroborates earlier studies that show that traditional assessment methods are the most used in higher education (Struyven et al., 2005; Flores et al., 2015; MacLellan, 2001). Although the test is the method most used, it is used differently depending on the field knowledge. In LHS all teachers used tests. In NES 8 out of 9 respondents used test such as in the case of ESE, in which 5 out of 6 teachers report using written tests. Regarding SSH teachers, 18 out of 32 interviewed teachers mentioned written tests as the most used assessment method. Furthermore, findings show that LHS, NES and ESE used more the test as the preferred method of assessment than their SSH counterparts (only over half of the teachers use it). However, other methods more focused on learner and skills' development are used by teachers as it is the case of practical work in groups that allows collaborative work or oral presentations that encourage students to develop communication skills.

Regarding the effectiveness of assessment methods, the participants state that the written test is the more effective assessment method. Other methods are also considered to be as effective methods such as group or individual work, methods based on continuous assessment, project-based work and oral presentations.

The participants point out that written tests require students' effort and they are more efficient than other assessment methods: "Exams require students to make an individual effort and that can be later seen as beneficial in their professional lives" (P.27 SSH); "Exams force students to make the information more systematic and they reflect they effort individually" (P.19 SSH).

However, others are critical of using written tests because they promote memorisation rather than understanding: "If you value memorisation, you choose the summative test, but students study and spend some time and after that they no longer know anything" (P.52 SSH), "I don't use tests and exams because for me they are episodes of memorisation" (P.34 SSH), "Methods that foster the memorisation rather than understanding such as tests do not enhance student learning" (P.46 LHS).

On the other hand, teachers claim that the assessment methods used on the basis of a continuous assessment are more effective.

Continuous assessment through inquiry-based work and project-based work foster a process of more continuous study, more discussion between students and teachers and amongst students and promote the development of soft skills. (P.13 ESE)

Formative feedback and systematic monitoring of learning are also the focus of continuous assessment in the participants' perspective.

The methods that enable continuous assessment are better. They enable the monitoring of individual work and self and peer assessment. Students can become more aware of their learning, overcome shortcomings, receive guidance, and be stimulated in order to feel more comfortable and safe in their own development and progression. (P.16 SSH)

Group work, oral presentations and projects are also pointed out by the participants. Group work is more effective because it promotes the development of communication skills, collaboration and it improves learning.

Group work is better because it implies more interaction between students and the teacher. It also leads to greater learning experience, not just memorisation and reproduction of knowledge. (P.54 SSH)

Oral presentations and projects are also identified as methods that promote research and self-monitoring of learning.

Projects and oral presentations of scientific papers are better methods to assess student learning. These methods provide students with opportunities to search for information and to self-regulate their learning. (P.26 ESE)

Although tests and exams were identified by teachers as the most effective methods to assess student learning, for some of them this method only promotes memorisation and it is used as a method for assessing large classes (Biggs, 2003; Pereira & Flores, 2012). Students' perceptions regarding traditional assessment methods are also negative because they think they are inadequate and superficial as a tool to measure learning (Struyven et al., 2005). However, it is recognised that traditional methods may be suitable for certain purposes and in given contexts of teaching and learning (Flores et al., 2015). On the other hand, there is evidence that methods such as group work, oral assessment and practical works have benefits for learning and are significant for students' professional life (MacLallen, 2004; Taylor, 1997). Group work is a good example, since it allows the development of communication skills, group management, dealing with problems (Almond, 2009; Johnston & Miles, 2004), and encourages students' motivation (Weurlander, Soderberg, Scheja, Hult, & Wernerson, 2012). However, Flores et al. (2015) study show that when students are assessed by the so-called learner-centred methods the idea of conflict is more associated with assessment than when other (more traditional) assessment methods are used.

5.3. Role of Assessment

Findings from this study suggest that the participants use *summative* and *formative assessment*, although summative assessment is more used. In teachers' accounts both positive and negative aspects in formative and summative assessment are identified.

Summative assessment is used by most of them because it is an institutional requirement of all Portuguese universities participating in this study: "I use summative assessment at the end of the semester because the assessment standards of the university require so" (P.4 SSH); "The final summative assessment is compulsory at my institution" (P.26 ESE).

Some teachers agree to use summative assessment since it is efficient and effective and can act in the end as "a safeguard for the teacher" (P.52 SSH).

However, most teachers look at summative assessment as an unfair process due to its mandatory and reductionist nature.

I use summative assessment because I have to. At the end of the day there is to be a mark and in this university all students are measured by a final grade, it is unfair. Sometimes students get 14.4 (out of 20) and will not have 15 and a student with 14.5 gets 15 at the end. (P.43 LHS)

Also, the difficulty of having to translate the entire assessment into a summative grade is reported by the participants: “I use summative assessment because I am forced to translate everything into a grade, because the university selects people” (P.47 SSH); “I use summative assessment and in the end the excel programme is in charge, but it is not a fair assessment” (P.39 LHS).

Formative assessment presents multiple benefits to the learning process. The monitoring of learning and the knowledge construction are positive aspects reported by the participants.

Formative assessment improves knowledge construction process and allows the reformulation of my own performance and students' performance. (P.25 SSH)

I think formative assessment is important. If I could I wouldn't give grades. For me it is more important to see what the students do, the difficulties they have ... and then having to turn it into numbers is really hard. (P.39 LHS)

Feedback is also a key feature which continually regulates the learning process, and it is recognised by the participants as a tool to improve the learning process: “The cornerstone and the basic assumption is formative assessment. The grades come after that and result from continuous feedback” (P.40 SSH). Also:

I use formative assessment. In every classroom we discussed what they learned and students have contact with what they are learning. We do exercises and I only ask the students what I am capable of assessing, so they feel that all I ask is useful. I think that feedback is crucial, so all I ask I have to have time to look and analyse. (P. 50 SSH)

The participants also highlight that formative assessment enables students to develop competencies and to improve their performance.

I see many advantages in formative assessment, because we cannot see the process as a product; there is also the development of skills and feedback to the students to improve their learning. It is important for them in order to change their behaviours. And there is only learning when there is change of a particular behaviour. (P.49 LHS)

In addition to improve students' performance, teachers hold the idea that formative assessment also allows the regulation of teaching and learning process: “I always use formative assessment. The idea is that the student in the end meets the goals and this requires reformulating the teaching and learning process” (P.24 SSH).

Other benefits were also associated with formative assessment such as encouragement of the critical thinking:

Formative assessment helps students to understand how they produced the materials, what kinds of mistakes they made and they may incorporate formative feedback in their next work, improving their own working style, intellectual study, learning, improving the methods for learning and developing their skills. For example, when it is noted that a student is more reproductive he/she is encouraged to take a more analytical and critical stance. (P. 41 SSH)

The fairness of formative assessment is also presented in the participants' accounts:

Formative assessment is fairer and better for the student. Because if the student only performs a test he/she can be in a bad day and something happens and will affect all the work that he/she developed until then. (P. 46 LHS)

There was also evidence of formative assessment as the assessment that is more suitable to higher education purposes.

Formative assessment is the most important. It corresponds to my idea of university and the role of the teacher and his/her work. It corresponds to my idea of ethics and responsibility to the student. It is a way of seeing education (P.47 SSH) or I use formative assessment because it is the assessment that best responds to the modernisation of higher education. (P.26 ESE)

Although formative assessment is said to have benefits for learning, some participants state that it is not feasible in higher education due to the number of students, lack of resources and available time: “We are unable to use formative assessment and formative assessment is important. In a lecture hall with 100 students it is unthinkable” (P.45 LHS); “In undergraduate education the formative assessment does not make sense, it only makes sense at the master degree level” (P.35 SSH). Also:

I cannot use formative assessment because I have no time and it forces me to do something that I'm not good at, I mean to distribute the interactivity with the students and it is difficult. (P.37 NES)

These findings are in line with Gibbs and Simpson (2004) who identified constraints that reduce the use of formative assessment in higher education: short courses, consequently less contact hours, increase in the number of examinations and issues related to staff. These constraints inhibit the quality and quantity of feedback given and influence the use of assessment practices.

Finally, some participants did not answer if they use or not formative and summative assessment because they did not know what it meant. Findings reveal that they pointed out more positive aspects related to formative assessment rather than to summative assessment (see figure 1). In the participants' opinion, the positive aspects of formative assessment are related to the teaching and learning process. These positive aspects reveal a concern with the learning process regarding knowledge, monitoring of learning, skills' development and feedback. Positive aspects are associated with teachers' conceptions on assessment: fairness and the idea of university. With regard to summative assessment, the positive aspects related to the rationality of effectiveness and efficacy of assessment process and as a proof that they can serve as justification if the teacher needs one. The constraints of formative assessment are related to lack of time, large number of the students, heavy workload whereas the constraints of summative assessment are associated with an institutional obligation and its unfairness nature. In general, formative assessment is related to the improvement of learning and requires certain conditions to be implemented in an effective manner. Summative assessment is associated with systematisation of assessment concerning efficacy and effectiveness, and as a teacher's safeguard as well as an institutional requirement.

Other studies corroborate these findings and show that assessment in higher education serves different purposes (Boud, 1995). MacLellan (2001) found that summative assessment is seen by both teachers and students as the purpose of assessment and synonymous with give and receive marks. The fact of assessment is seen as the purpose of assessment may be related to the summative assessment in some contexts as an institutional requirement. On the other hand, the study by Samuelowicz and Bain (2002) concerning teachers' orientations to assessment practice shows that teachers view the purpose of assessment as support to the students' learning and feedback.

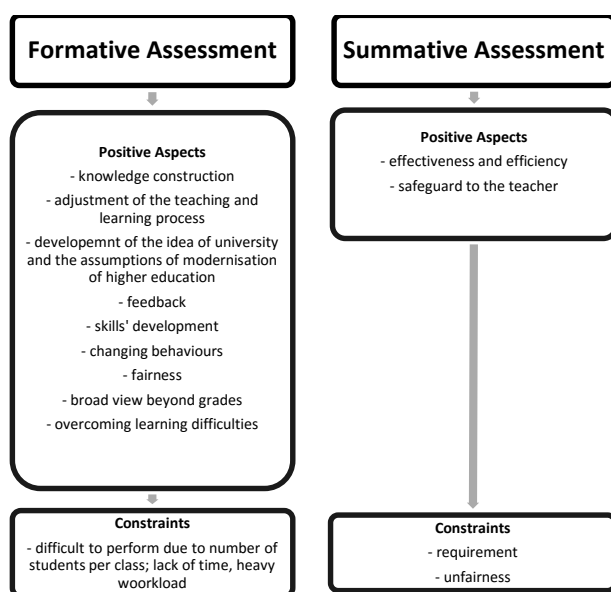


Fig 1. Positive aspects and constraints of summative and formative assessment
Source: Prepared by the authors.

5.4. Key moments in which assessment is used

Teachers assess students' learning in three different moments: (1) during the process of teaching and learning and at the end of the course; 2) during the process of teaching; and 3) every time students perform a task.

Most of the teachers assess students' learning during the process of teaching and learning and at the end of the course: "I assess during the semester, after the delivery of students' work. At the end of the semester after students' perform an exam" (P.13 ESE); "I assess during the semester and at the end. However, I assess during the semester only when there is sufficient topics to be reviewed" (P.27 SSH); "There are always several moments of assessment during the semester and final examinations must be offered to the students according to the university Regulation" (P.26 ESE).

The phase of assessing during the process is especially related to continuous assessment: "As I use continuous assessment, it is being done throughout the semester and at the end there is a more formal self and peer assessment" (P.25 SSH); "I use formative assessment during the semester and summative assessment at the end with group work" (P.22 SSH).

Other participants state that the assessment is carried out only during the semester and again the idea of a continuous assessment is highlighted. The participants claim that continuous assessment brings benefits for students and it is more effective:

It makes sense to develop continuous assessment throughout the semester, it fosters students' motivation to come to classes and it is effective and continuous. If they are missing they are not admitted to the exam. (P.42 ESE)

I use continuous assessment. It is the commitment between having a perspective of what will happen and the feasibility of what is possible because to assess all classes is impossible. (P.55 SSH)

However, the participants also explain that there are certain rules that do not allow them to perform the assessment only during the process: "I assess my students throughout the Curricular Unit. Then there are formal moments due to the constraints

that we have” (P. 43 LHS); “The Pedagogical Council determines the periods of written assessments (test/exam). The remaining assessments occur during the semester” (P.8 SSH).

Some accounts also reveal that the participants assess students’ learning every time a student performs a task: “I always try to provide students with feedback when they perform a task, an activity. Because I believe that this is the most effective way to assess” (P.16 SSH); “I assess in all classes, whenever they do a task” (P.29 SSH).

From other university teachers’ accounts assessment is also carried out only at the end of the course: “As we have few weeks of classes I only assess at the end of the assessment elements” (P.45LHS), “I assess only at the end of the year” (P.30 SSH).

Also a periodic assessment is reported by the participants: “The assessment of attitudes is continuous and the assessment of knowledge acquired is done in given moments” (P. 48 LHS).

Some participants explain that they assess all the time: “I assess all the time to maintain self-learning and self-regulation” (P.4 SSH); “I assess throughout the semester, but especially over the last month, when the students have assimilated much of the content” (P.2 SSH).

Other university teachers, however, claim that it: “depends on the nature of the Curricular Unit” (P. 1 ESE).

In the participants’ accounts it is clear the idea of the importance of continuous assessment throughout the process with benefits for the learning process. However, when most teachers assess during the semester, they normally use tests and they associate them with university requirements. A number of constrains were also identified in their accounts, namely the imposition of assessment. The idea of assessing only throughout the process of teaching and learning allows them a given effectiveness in assessment as it requires students to attend classes. Other participants claim that they assess every time a student performs a task emphasising the importance of timely and continuous feedback. On the other hand, none of participants mentioned that assessing at the beginning of the semester or at the beginning of a module. The participants seem to use formative assessment in their practices, however, they do not assess at the beginning of a module or coursework. One might question if they actually use formative assessment and how they use it and what kind of feedback they provide and when as although teachers say they use formative assessment, their practices go against it (MacLellan, 2001).

6. Discussion

This paper sets out to analyse the conceptions of assessment and assessment methods used in higher education from teachers’ point of view. Data highlight that assessment that can help students to learn better is continuous assessment. The participants view formative assessment as important, although most of them use summative assessment because it is compulsory in higher education institutions. In their perspective, formative assessment improves the learning process, allows the monitoring of learning and feedback, identifies learning difficulties and it is a fairer assessment mode. However, together with continuous assessment, formative assessment is seen as not feasible to implement in higher education due to the number of students per class, lack of resources, heavy workload and lack of time. Given these constraints, teachers resist

to use certain forms of assessment such as continuous, formative assessment and the use of participatory methods. On the other hand, summative assessment is seen by university teachers as more efficient and effective, as it can work as a proof of student performance within the institution. However, summative assessment is also regarded by them as compulsory and unfair. Although some participants claim that they use formative assessment, one might question its purpose and effect taken into account the moments of assessment and the nature of the methods that are used.

The participants look at assessment as a continuous process throughout teaching and learning as it brings benefits such as learning monitoring through feedback and students' attendance to classes. Although they state that continuous assessment is important, they also recognise that there must be final formal assessment methods, which are determined by the university, in some cases, and normally it is a written test at the end of the semester, which they see as a formal constraint.

Thus the participants hold different conceptions of assessment and one might question how continuous and formative assessment is understood. The participants tend to associate more continuous assessment with written tests throughout the semester and less with formative and timely feedback to students.

Some participants see assessment as summative emphasising its effective dimension and reporting its mandatory nature. Other teachers see assessment as formative as synonymous with development, learning, knowledge, adjustment and improvement. The divergence of perspectives is associated with the lack of clear institutional policy on assessment practices and previous experiences of the university teachers (Fletcher et al., 2012; McMillan, 2003). As a result, research is needed about how conceptions of assessment can be influenced and enhanced the institutional policy (Rust, 2007; Brown, 2004).

The written test is the most used assessment method. Practical work in groups, oral presentations, individual work and reports are also used too but their use is less frequent. Teachers recognise the importance of assessment methods that allow continuous assessment, guidance, monitoring of learning and skills' development. Some of them also recognise that traditional methods promote memorisation rather than knowledge. Although they recognise the importance of student-centred assessment methods, written tests continue to be the most used method identified by the participants. As Ramsden (2004) states, there is not a method that satisfies all educational goals, therefore diversity and a balance in assessment methods are key issues in order to innovate and improve assessment practices.

7. Conclusion

In conclusion, this study provides evidence that there is a contradiction between the assessment practices that the participants claim to be most beneficial to students' learning and those they claim to use in practice. This contradiction between their conceptions of assessment and the practices they claim to use may be explained by different reasons: heavy workload, lack of human and physical resources in higher education contexts and the obligation of using summative assessment imposed by universities which inhibits the use of practices centred on the learner perpetuating, thus, the use of traditional assessment methods. The results of this study will contribute to understand how assessment process is carried out in some contexts within the Portuguese higher education.

This study suggests recommendations for further research. It would be important to understand if the university teachers' conceptions of assessment have direct influence on their assessment practices. It would also be important to understand why teachers' conceptions of assessment are different from those that are put in practice by them as well as the perceptions of their students. Also, further research is needed on university teachers' conceptions of assessment and their relation with students' achievement (Brown & Hirschfeld, 2008).

References

- Almond, R. (2009). Group assessment: Comparing group and individual undergraduate module marks. *Assessment & Evaluation in Higher Education*, 34(2), 141-148. doi:10.1080/02602930801956083.
- Atkins, M. (2004). What should we be assessing? In P. Knight (Ed.), *Assessment for learning in higher education* (pp. 25-33). Birmingham: SEDA.
- Biggs, J. (2003). *Teaching for quality learning at university*. Buckingham: SHRE & Open University Press.
- Birenbaum, M., & Feldman, R. (1998). Relationships between learning patterns and attitudes towards two assessment formats. *Educational Research*, 40(1), 90-98. doi: 10.1080/0013188980400109.
- Black, P., & Wiliam, D. (1998). *Inside the Black Box: Raising standards through classroom assessment*. London: School of Education, King's College.
- Blaikie, N. (2010). *Designing social research*. United Kingdom: Polity Press.
- Bloom, B., Hastings, J., & Madaus, G. (1971). *Manual de avaliação formativa e somativa do aprendizado escolar*. São Paulo: Pioneira.
- Boud, D. (1995). Assessment and Learning: contradictory or complementary? In P. Knight (ed.), *Assessment for learning in higher education* (pp. 35-48). London: Kogan Page Limited.
- Boud, D. (2000) Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 151-167. doi:10.1080/713695728.
- Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education*, 31(4), 399-413. doi:10.1080/02602930600679050
- Boud, D., & Falchikov, N. (2007). *Rethinking assessment in higher education: Learning for the long term*. New York: Routledge.
- Brew, C., Riley, P., & Walta, C. (2009). Education students and their teachers: Comparing views on participative assessment practices. *Assessment and Evaluation in Higher Education*, 34(6), 641-657. doi:10.1080/02602930802468567.
- Brown, G. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. *Assessment in Education: Principles, Policy & Practice*, 11(3), 301-318. doi:10.1080/0969594042000304609
- Brown, G., Bull, J., & Pendlebury, M. (1997). *Assessing student learning in higher education*. London: Routledge.
- Brown, G., & Hirschfeld, G. (2008). Students' conceptions of assessment: Links to outcomes. *Assessment in Education: Principles, Policy & Practice*, 15(1), 3-17. doi:10.1080/09695940701876003.
- Brown, S., & Knight, P. (1994). *Assessing learners in higher education*. London: Kogan Page.

- Brown, G., Lake, R., & Matters, G. (2011). Queensland teachers' conceptions of assessment: The impact of policy priorities on teacher attitudes. *Teaching and Teacher Education, 27*(1), 210-220. doi:10.1016/j.tate.2010.08.003.
- Brown, S., Race, P., & Rust, C. (2004). Using and experiencing assessment. In P. Knight (Ed.), *Assessment for learning in higher education* (pp. 75-85). Birmingham: SEDA.
- Bucharest Communiqué. (2012). *Making the most of our potential: Consolidating the European Higher Education Area*.
- Cohen, L., Manion, L., & Morrison, K. (2008). *Research methods in education*. London: Routledge.
- Dochy, F., & McDowell, L. (1997). Assessment as a tool for learning. *Studies in Educational Evaluation, 23*(4), 279-298. doi:10.1016/S0191-491X(97)86211-6.
- Dochy, F., Segers, M., Gijbels, D., & Struyven, K. (2007). Assessment engineering: Breaking down barriers between teaching, and learning and assessment. In D. Boud & N. Falchikov (Eds.), *Rethinking assessment in higher education: Learning for the longer term* (pp. 87-100). New York: Routledge.
- Drew, S. (2001). Student perceptions of what helps them learn and develop. *Teaching in Higher Education, 6*(3), 309-331. doi:10.1080/13562510120061197.
- Fernandes, S., Flores, M., & Lima, R. (2012). Students' views of assessment in project-led engineering education: Findings from a case study in Portugal. *Assessment & Evaluation in Higher Education, 37*(2), 163-178. doi:10.1080/02602938.2010.515015.
- Fletcher, R., Meyer, L., Anderson, H., Johnston, P., & Rees, M. (2012). Faculty and students conceptions of assessment in higher education. *Higher Education, 64*(1), 119- 133. doi:10.1007/s10734-011-9484-1.
- Flores, M., & Veiga Simão, A. M. (2007). Competências desenvolvidas no contexto do Ensino Superior: A perspectiva dos diplomados. In M. Iglesias Martínez & F. Pastor Verdú (Eds.), *Actas das V Jornadas de Redes de Investigación en Docencia Universitaria*. Alicante: Universidad de Alicante.
- Flores, M., Veiga Simão, M., Barros, A., & Pereira, D. (2015). Perceptions of effectiveness, fairness and feedback of assessment methods: A study in higher education. *Studies in Higher Education, 40*(9), 1523-1534. doi:10.1080/03075079.2014.881348.
- Gibbs, G. (1999). Using assessment strategically to change the way students learn. In S. Brown & A. Glasner (Eds.), *Assessment matters in higher education: Choosing and using diverse approaches* (pp. 41-53). Buckingham: SHRE & Open University Press.
- Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education, 1*, 3-31. Retrieved from <http://www.open.ac.uk/fast/pdfs/Gibbs%20and%20Simpson%202004-05.pdf>
- Hadji, C. (1994). *A Avaliação, regras do jogo: Das intenções aos instrumentos*. Porto: Porto Editora.
- Hue, M., Leung, C., & Kennedy, K. (2014). Student perception of assessment practices: Towards 'no loser' classrooms for all students in the ethnic minority schools in Hong Kong. *Educational Assessment Evaluation and Accountability, 27*(3), 253-273. doi:10.1007/s11092-014-9205-2.
- Johnston, L., & Miles, L. (2004). Assessing contributions to group assignments. *Assessment & Evaluation in Higher Education, 29*(6), 751-768. doi:10.1080/0260293042000227272.
- Kvale, S. (1996). *Interviews: An introduction to qualitative research interviewing*. London: SAGE Publications.
- Leuven/Louvain-la-Neuve Communiqué. (2009). *Communiqué of the Conference of European Ministers Responsible for Higher Education*.

- Light, G., & Cox, R. (2003). *Learning and teaching in higher education: The reflective professional*. London: Sage Publications.
- MacLellan, E. (2001). Assessment for learning: The differing perceptions of tutors and students. *Assessment & Evaluation in Higher Education*, 26(4), 307-318. doi:10.1080/02602930120063466.
- MacLellan, E. (2004). Authenticity in assessment tasks: A heuristic exploration of academics' perceptions. *Higher Education Research and Development*, 23(1), 19-33. doi:10.1080/0729436032000168478.
- Marton, F., & Saljo, R. (1997). Approaches to learning. In F. Marton, D. Hounsell, & N. Entwistle (Eds.), *The experience of learning. Implications for teaching and studying in higher education* (pp. 39-58). Edinburgh: Scottish Academic Press.
- McMillan, J. H. (2003). Understanding and improving teachers' classroom assessment decision making: Implications for theory and practice. *Educational Measurement: Issues and Practice*, 22(4), 34-43. doi:10.1111/j.1745-3992.2003.tb00142.x.
- Meyers, N., & 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. doi:10.1111/j.1745-3992.2003.tb00142.x.
- Miles, M., & Huberman, M. (1994). *Qualitative data analysis* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Miles, M., Huberman, M., & Saldaña, J. (2014). *Qualitative data analysis. A methods sourcebook*. Thousand Oaks, CA: Sage Publications.
- Pajares, F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62(3), 307-332. doi:10.3102/00346543062003307.
- Pereira, D., Flores, M., & Niklasson, L. (2015). Assessment revisited: A review of research in *Assessment and Evaluation in Higher Education*. *Assessment & Evaluation in Higher Education*. doi:10.1080/02602938.2015.1055233.
- Pereira, D., & Flores, M. (2012). Percepções dos estudantes universitários sobre a avaliação das aprendizagens: Um estudo exploratório. *Avaliação (Campinas)*, 17(2), 529-56. doi:10.1590/S1414-40772012000200012.
- Perrenoud, P. (1999). *Avaliação, Da excelência à regulação das aprendizagens: Entre duas lógicas*. Porto Alegre: ARTMED.
- Portela, M., Sá, C., Alexandre, F., & Cardoso, A. (2009). Perceptions of the Bologna Process: What do students' choices reveal? *Higher Education*, 58(4), 465-474. doi:10.1007/s10734-009-9205-1.
- Postareff, L., & Lindblom-Ylänne, S. (2008). Variation in teachers' descriptions of teaching: Broadening the understanding of teaching in higher education. *Learning and Instruction*, 18(2), 109-120. doi:10.1016/j.learninstruc.2007.01.008.
- Postareff, L., Virtanen, V., Katajavuori, N., & Lindblom-Ylänne, S. (2012). Academics' conceptions of assessment and their assessment practices. *Studies in Educational Evaluation*, 38(3-4), 84-92. doi:10.1016/j.stueduc.2012.06.003.
- Prosser, M., & Trigwell, K. (1998). *Understanding leaning and teaching: The experience in higher education*. Buckingham: SHRE & Open University Press.
- Ramsden, P. (2004). *Learning to teach in higher education*. London: Routledge Falmer.

- Remesal, A. (2011). Primary and secondary teachers' conceptions of assessment: A qualitative study. *Teaching and Teacher Education*, 27(2), 472-482. doi:10.1016/j.tate.2010.09.017.
- Rowntree, D. (1987). *Assessing students: How shall we know them?* London: Kogan Page.
- Rust, C. (2007). Towards a scholarship of assessment. *Assessment & Evaluation in Higher Education*, 32(2), 229-237. doi:10.1080/02602930600805192.
- Sambell, K., & McDowell, L. (1998). The values of self and peer assessment to the developing lifelong learner. In C. Rust (Ed.), *Improving student learning: Improving students as learners* (pp.56-66). Oxford: Oxford Center for Staff and Learning Development.
- Samuelowicz, K., & Bain, J. D. (2002) Identifying academics' orientations to assessment practice. *Higher Education*, 43(2), 173-201. doi:10.1023/A:1013796916022.
- Schreier, M. (2012). *Qualitative content analysis in practice*. London: SAGE Publications.
- Scouller, K. (1998). The influence of assessment method on students' learning approaches: Multiple choice question examinations versus assignment essay. *Higher Education*, 35(4), 453-472. doi:10.1023/A:1003196224280.
- Segers, M., Gijbels, D., & Thurlings, M. (2008). The relationship between students' perceptions of portfolio assessment practice and their approaches to learning. *Educational Studies*, 34(1), 35 -44. doi:10.1080/03055690701785269.
- Silverman, D. (2013). *Doing qualitative research: A practical handbook*. London: SAGE Publications.
- Simão, J. V., Santos, S., & Costa, A. (2003). *Ensino Superior: Uma visão para a próxima década*. Lisbon: Gradiva.
- Simms, M., & George, B. (2014). Approaching assessment from a learning perspective: Elevating assessment beyond technique. *Educational Assessment Evaluation and Accountability*, 26(1), 95- 104. doi:10.1007/s11092-013-9176-8.
- Sluijsmans, D., Dochy, F., & Moerkerke, G. (1998). Creating a learning environment by using self, peer and co-assessment. *Learning Environment Research*, 1(3), 293-319. doi:10.1023/A:1009932704458.
- Struyven, K., Dochy, F., & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: A review. *Assessment & Evaluation in Higher Education*, 30(4), 331-347. doi:10.1080/02602930500099102.
- Tang, C., Lai, P., Arthur, D., & Leung, S. (1999). How do students prepare for traditional and portfolio assessment in a problem-based learning curriculum? In J Conway & A. Williams (Eds.), *Themes and Variations in PBL: Refereed proceedings of the 1999 Bi-ennial PBL Conference Vol.1* (pp. 206-217). Sydney: Australia Problem-Based Learning Network (PROBLARC).
- Taylor, I. (1997). *Developing learning in professional education*. Buckingham: SHRE & The Open University Press.
- Thomas, R. (2003). *Blending qualitative and quantitative research methods in theses and dissertations*. Thousand Oaks, CA: Sage Publications.
- Thompson, A. G. (1992). Teachers' beliefs and conceptions: A synthesis of the research. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 127-146). New York: Macmillan.
- Trigwell, K., & Prosser, M. (1996). Changing approaches to teaching: A relational perspective. *Studies in Higher Education*, 21(3), 275-284. doi:10.1080/03075079612331381211.

- Watering, G., Gijbels, D., Dochy, F., & Rijt, J. (2008). Students' assessment preferences, perceptions of assessment and their relationships to study results. *Higher Education*, 56(6), 645-658. doi:10.1007/s10734-008-9116-6.
- Watkins, D., Dahlin, B., & Ekholm, M. (2005). Awareness of the backwash effect of assessment: A phenomenographic study of the views of Hong Kong and Swedish lecturers. *Instructional Science*, 33(4), 283-309. doi:10.1007/s11251-005-3002-8.
- Weber, R. (1990). *Basic content analysis*. Newbury Park, CA: Sage Publications.
- Webber, K. (2012). The use of learner-centered assessment in US colleges and universities. *Research in Higher Education*, 53(2), 201-228. doi:10.1007/s11162-011-9245-0.
- Weurlander, M., Soderberg, M., Scheja, M., Hult, H., & Wernerson, A. (2012). Exploring formative assessment as a tool for learning: Students' experiences of different methods of formative assessment. *Assessment & Evaluation in Higher Education*, 37(6), 747-760. doi:10.1080/02602938.2011.572153.

Appendix 1. Interview schedule (example)

Questions related to conceptions of assessment

In your opinion, taking into account your experience, what kind of assessment can help students learn better?

Do you use formative assessment? Why?

And summative assessment? Why and when?

In general, when you assess the students' learning? At the beginning of the semester, during the semester, at the end of the semester, or whenever the student performs a task? Why?

Questions related to assessment methods

What are the assessment methods you use the most?

Taking into account your experience as a teacher, what are the most effective methods of assessment in terms of teaching and learning? Why?

La Evaluación a través de Portafolios: ¿Una Ocasión para el Aprendizaje?

The Assessment Through Portfolios: An Opportunity for Learning?

Noelia Alcaraz Salarirche*

Universidad de Cádiz

Autores como Calfee y Perfumo (1993), Shores y Grace (2004), Delamstro (2005), Barberá (2005), Darling-Hammond (2005) o Pérez Rendón (2014) defienden un uso del portafolios centrado en el aprendizaje del-la estudiante y resaltan la vertiente reflexiva del mismo como una de las condiciones claves de la enseñanza y la evaluación a través de portafolios.

En este trabajo se exponen aquellas otras condiciones educativas –relacionadas con la reflexión sobre el aprendizaje y la enseñanza– que se han de tener en cuenta al poner en marcha un sistema de evaluación a través de portafolios.

Antes, se hace también un breve recorrido por el origen del portafolios, así como por los principales tipos, usos y definiciones que se encuentran del mismo.

El objetivo de este artículo es situar al lector ante un nuevo escenario de evaluación y brindar posibilidades alternativas de pensar la evaluación en un aula. Evaluación no confundida con la calificación, sino como proceso de comprensión permanente que busca la mejora tanto de la enseñanza como del aprendizaje.

Palabras claves: Evaluación, Portafolios, Reflexión, Retroalimentación, Enseñanza, Aprendizaje.

Authors like Calfee and Perfumo (1993), Shores and Grace (2004), Delamstro (2005), Barber (2005), Darling-Hammond (2005), or Pérez Rendón (2014) advocate a use of portfolio focused on the students' learning and highlight the reflective side of it as one of the key conditions of teaching and assessment through portfolios.

This paper shows those other educational conditions –related to reflection on learning and teaching– which should be taken into account when implementing a system of assessment through portfolios.

Before, the work provides an overview of the origin of the portfolio, as well as its main types, uses and definitions.

The aim of this article is to place the reader in a new assessment scenario and provide alternative possibilities of thinking about classroom assessment. This assessment should not be confused with a grade (US) / mark (UK), but seen as a permanent understanding process that seeks to improve both teaching and learning.

Keywords: Evaluation, Assessment, Portfolio, Reflection, Feedback, Teaching, Learning.

*Contacto: noelia.alcaraz@uca.es

issn: 1989-0397

www.rinace.net/riee/

<https://revistas.uam.es/riee>

Recibido: 24 de septiembre de 2015

1ª Evaluación: 25 de octubre de 2015

Aceptado: 1 de diciembre de 2015

1. Introducción¹

El portafolios como herramienta para evaluar el proceso de aprendizaje del alumnado que cursa la etapa obligatoria de la enseñanza es una cuestión que resulta relativamente novedosa. Ello se debe a su escasa presencia y corta tradición en el ámbito educativo ya que este cuenta con una mayor trayectoria en contextos profesionales como los artísticos. Si bien es cierto que el uso del portafolios ha proliferado especialmente en los niveles de educación superior, el portafolios en su sentido literal, de ‘portar folios’, no es tan reciente. Lo novedoso del asunto sería otra cuestión: el abandono literal de la palabra, que convierte al portafolios en una práctica educativa al servicio de la evaluación y del aprendizaje. Es por ello que, como veremos a lo largo de este artículo, para que el portafolios sea algo más que un contenedor de trabajos de estudiantes que sirven para mostrar y demostrar las capacidades o competencias que estos poseen, se deben establecer una serie de condiciones que las prácticas con portafolios deben cumplir, siempre que estas pretendan estar al servicio de la evaluación –evaluación en el sentido de búsqueda de la mejora–. Es el reto ante el que nos encontramos: apostar por el diseño de un nuevo sistema de evaluación basado en la reflexión, la comunicación, la comprensión del error y el análisis de los procesos y contextos de aprendizaje. Se trata de una oportunidad para aprovechar la escasa tradición de prácticas con portafolios en las escuelas, y el carácter continuo y reflexivo que se le otorga en la literatura, para alejarnos del uso acreditativo y selectivo que este ha tenido en otros ámbitos.

2. Breve recorrido por el origen, los tipos y los usos del portafolios

2.1. *El nacimiento del portafolios en el ámbito educativo*

Históricamente el uso de los portafolios se sitúa en contextos profesionales relacionados sobre todo con las artes, la publicidad o las finanzas. Así lo exponen Danyelson y Abrutyn (1999): “Los portafolios, desde luego, no son una invención reciente. El concepto existe desde hace mucho en numerosos ámbitos fuera del aula. Artistas, arquitectos y fotógrafos lo usan para presentar su trabajo a clientes potenciales” (pp. 7-8).

En estos contextos los portafolios profesionales tenían y tienen como finalidad mostrar las mejores producciones elaboradas a lo largo de una determinada trayectoria profesional, debiendo cumplir con las demandas de intereses de los clientes a los que han de ser mostrados.

En palabras de Oosterhof (1996):

En esas profesiones, una persona compone un portafolios que contiene ejemplos de trabajo que se muestran a los clientes potenciales y a los empleadores. Los portafolios incluyen

¹ El contenido del presente artículo forma parte de la Tesis Doctoral de la autora, dirigida por Ángel I. Pérez Gómez, en el Departamento de Didáctica y Organización Escolar de la Facultad de Ciencias de la Educación, de la Universidad de Málaga.

generalmente el mejor trabajo. Por otra parte, el contenido del portafolios también debe responder a los intereses y necesidades del cliente o empleador. (p. 165)

En este sentido, el uso que desde varios campos profesionales se ha venido dando al portafolios hace que este se presente como la vía a través de la cual profesionales o futuros profesionales debían demostrar sus conocimientos, habilidades y competencias en un determinado ámbito profesional, para ser valorado por el destinatario de dicho portafolios.

Es en la década de los noventa cuando en Europa empieza a extenderse hacia el ámbito educativo el interés por el portafolios. Los noventa, sobre todo a partir de su tercera década, se caracterizan, de acuerdo con Alcaraz (2015a), por la aparición de las ideas de Tyler, cuya intención era superar la mera evaluación psicológica. Es decir, el interés por el portafolios en el campo de la educación se sitúa especialmente en EE. UU. y Reino Unido, cuando empieza a surgir cierto descontento con las pruebas estandarizadas tipo test que tan a la orden del día estaban.

Este desarrollo se produce principalmente en el ámbito de la lectoescritura (en su estudio como lengua materna o como segunda lengua). Según Oosterhof (1996), una de las ventajas del portafolios con respecto a lo que hasta el momento existía –refiriéndose a las tradicionales pruebas de desempeño académico– era su fácil adaptabilidad a las metas educativas de cada estudiante. Sostenía que las pruebas escritas a menudo implicaban un conjunto común de ítems administrados simultáneamente a un grupo de estudiantes, mientras que “los portafolios, sin embargo, pueden ser totalmente individualizados. Cada portafolios puede ser sensible a los objetivos establecidos por cada estudiante” (Oosterhof, 1996, p. 166).

Por su parte, Delmastro (2005), que también sitúa el desarrollo histórico del portafolios y su mayor proliferación hacia mediados de los noventa en EE. UU. y Canadá, sostiene que es en los setenta cuando la UNESCO, con sus programas a nivel internacional para el desarrollo de la lectoescritura, comienza a provocar un mayor auge y difusión del portafolios como estrategia de evaluación en países del hemisferio occidental. La autora recoge que la mayoría de las investigaciones realizadas en ese ámbito contemplan la enseñanza y evaluación de destrezas de lectoescritura en el contexto del aprendizaje y desarrollo de la lengua materna (Hamilton, 1994; Jalbert, 1998; Murphy y Camp, 1996; Wile y Tierney, 1996; citados por Delmastro, 2005). Así mismo, Bastidas (1996), Reckase (1995) y Delmastro (2005) exponen que el uso del portafolios se ha ido extendiendo hacia Israel, Francia, Italia y Australia, lográndose avances en la incorporación del portafolios en educación desde el ámbito institucional, incluso llegándose a explorar sus posibilidades en la evaluación a gran escala en instituciones de EE. UU. y Canadá.

En este sentido, French (1992) centra la atención en el portafolios utilizado para atender a estudiantes de grupos minoritarios que aprenden el inglés como segunda lengua, que no son valorados de manera correcta por pruebas estandarizadas que no responden a sus características específicas.

Por otro lado, Lyons (1999) sitúa el origen del portafolios en la insatisfacción que existía con respecto a las escuelas norteamericanas en las décadas de los setenta y ochenta. De modo que la experimentación con portafolios parecía responder a la necesidad de implementar innovaciones focalizando la atención en docentes y estudiantes, inicialmente en relación con la evaluación y de manera más reciente vinculada a la formación de docentes.

La consecuencia de todo ello es la proliferación en EE. UU. de multitud de experiencias escolares que introducen la práctica del portafolios como herramienta alternativa de evaluación. Surgen estudios, investigaciones y publicaciones sobre el tema en cuestión que tratan de documentar los procesos de implementación del portafolios en la enseñanza obligatoria. Rober C. Calfee y Pam Perfumo publican en 1993 en una revista llamada *Journal of Reading* un artículo titulado “Student Portfolios: Opportunities for a Revolution in Assessment” (“El Portafolios de Estudiante: Oportunidades Para una Revolución en Evaluación”). En él realizan un análisis del portafolios a partir de una investigación en la que diferentes centros hacen uso de este nuevo sistema de carpetas. Esto provocó que especialistas en psicometría se mostrasen inquietos ante tales acontecimientos, ya que la principal preocupación se centraba en cómo arbitrar mecanismos que permitiesen la estandarización del rendimiento y que cumpliesen con criterios de fiabilidad y validez. Sin embargo, eran los propios enseñantes quienes empezaban a tomar la iniciativa en el movimiento de sustitución de las pruebas estandarizadas (en parte o en su totalidad) por portafolios (Pelavin, 1991, citado por Calfee y Perfumo, 1993).

En España, no es hasta finales del siglo XX y principios del XXI cuando se empiezan a vislumbrar experiencias escolares que ponen en práctica el uso del portafolios. Contamos por tanto con una corta experiencia en el desarrollo de estas, especialmente en la educación obligatoria. Sin embargo, es cada vez más frecuente encontrarnos con experiencias de portafolios o de e-portafolios, sobre todo en el ámbito de la educación superior, de modo que la literatura educativa española que versa sobre el portafolios también comienza a proliferar.

Así por ejemplo, para Barberá (2005), el interés por el portafolios surge de un creciente interés por la elaboración de instrumentos y materiales útiles, acordes a los nuevos retos educativos. Según Wiggins (1990), la aparición de esta nueva práctica es un indicador de la búsqueda de mejoras en el contexto educativo, pretendiendo subsanar algunas importantes deficiencias en los sistemas educativos tradicionales más inmovilistas (citado por Barberá, 2005). Otros autores y autoras como Darling-Hammond (2005), Díaz Barriga (2006), Monereo (2009) y Díaz Barriga y Pérez (2010) sostienen que la revisión de los portafolios, en los niveles superiores, se centra específicamente en los portafolios como instrumentos de “evaluación auténtica” de las competencias de los-as docentes. Díaz Barriga y Pérez (2010, p. 7) reconocen que “la entrada del discurso de las competencias en educación también ha resultado un factor clave que ha impulsado el interés por la evaluación mediante portafolios” (citado por Pérez Rendón, 2014, p. 26).

2.2. Definiciones y tipos de portafolios

Hace décadas que maestros y maestras incluyen en la lista del material escolar que se envía a las familias la conocida carpeta con gomillas. La idea, especialmente en los cursos inferiores, era que el niño o la niña incluyera en su carpeta cada uno de los trabajos que se iban realizando durante los tres trimestres del curso escolar para ser exhibidos a los familiares.

También en cursos superiores se ha conocido su existencia. La “carpeta”, “dossier”, “carpeta memoria” o “portafolios” se ha venido usando como contenedor-organizador del material trabajado por el alumnado.

En este sentido, el portafolios ya existía y era entendido como un archivador que contenía el material escolar del alumnado. Se trata, por tanto, de una concepción

completamente estática del portafolios que, poco a poco, va tratando de superarse en un intento por abandonar el sentido literal de este, como mero contenedor de los materiales producidos. Empieza a introducirse un nuevo componente que se convierte en uno de los elementos centrales de la literatura del portafolios: la implicación activa del alumnado en su propio proceso de aprendizaje a través de la reflexión conjunta con el profesor o la profesora.

Empieza a cuestionarse el sentido del portafolios como una acumulación simple de trabajos (García Hernández, 2000), sin propósitos referidos a la mejora del aprendizaje o de la práctica docente. En la formación inicial de profesorado cobra especial protagonismo en el prácticum, sirviendo a la reconstrucción y reelaboración del conocimiento práctico (Hidalgo y De la Blanca, 2013). Es decir, el término portafolios abandona su significado literal y comienza a asociarse a una serie de connotaciones más educativas relacionadas con los procesos de reflexión, de aprendizaje, de retroalimentación, de colaboración, de participación y de diálogo, entre otros, incorporándose así al lenguaje educativo.

Así, para Pollari (2000), se trata de un portafolios “cuando el alumno participa en ello en lugar de ser objeto evaluación” (Paulson, Paulson y Meyer, 1991, citado por Pollari, 2000, p. 18). El autor resalta el sentido del portafolios en tanto en cuanto favorece la autogestión y autorregulación del alumnado.

Un portafolios cuenta una historia. Es la historia de conocimiento. Conocimiento sobre cosas. Conocimiento sobre uno mismo. Conocimiento sobre la audiencia. Los portafolios son las propias historias de los estudiantes, de lo que ellos saben, del por qué ellos creen que lo saben y por qué otros podrían ser de la misma opinión. Un portafolios es una opinión construida. Los estudiantes demuestran lo que conocen con ejemplos de su trabajo. (Paulson y Paulson, 1991, p. 2)

Shores y Grace (2004) resaltan como objetivo último de la aplicación de portafolios el apoyo a la educación centrada en el alumnado. “El portafolios puede convertirse en el contexto en el que el niño o la niña reflexione sobre las ideas o conocimientos [...]. Además, pueden reflexionar sobre su progresión” (2004, pp. 29-30).

Para Villalobos, “la reflexión es lo que diferencia a los portafolios de las carpetas sobre escritura [...]. En parte es a través de la reflexión que las carpetas o recopilaciones se transforman en portafolios” (2002, p. 391).

Por su parte, Pérez Gómez entiende que “el portafolios se convierte en un punto de apoyo fundamental para la autoevaluación” (2012, p. 238). En este sentido coinciden Calero y Calero (2008), que fundamentan el uso del portafolios sobre la base del constructivismo. Para ellos, este puede llegar a ser un poderoso y flexible instrumento que está:

Adquiriendo una relevancia pedagógica importante como estrategia de aprendizaje y evaluación cooperativa, que guía al estudiante hacia la reflexión sobre sus propios aprendizajes y que le implica en el desarrollo de su capacidad de construir conocimiento sobre lo que aprende y cómo lo aprende. (2008, p. 19)

En definitiva, estas y otras definiciones parecidas muestran cómo el portafolios desde su aparición hasta la actualidad ha ido transformando su significado, despegándose cada vez más del sentido literal que tenía hasta ir incorporando un significado con connotaciones mucho más educativas. No obstante, el sentido y significado de un portafolios concreto dependerá del tipo de portafolios que estemos implementando, lo cual a su vez estará condicionado por el uso que la persona que lo pone en marcha pretende hacer de él.

Si se hace un recorrido por la literatura que versa sobre portafolios, es fácil apreciar que existen variados tipos de portafolios. Por ejemplo, en el ámbito universitario, donde tienen una mayor presencia, se pueden encontrar prácticas con portafolios en diferentes ramas del saber como son las ciencias de la salud, ciencias experimentales, ingenierías, artes, humanidades y ciencias sociales, unas con mayor carga pedagógica y otras con menor. De modo que la variedad es amplia. Recojamos de Meeus y Lois (2006), citados por Pérez, Alcaraz y Fernández (2016), algunas de las posibles clasificaciones de portafolios: portafolios de presentación, portafolios de trabajo, portafolios para la retroalimentación, portafolios de rendición de cuentas, portafolios de producto, portafolios de evaluación, portafolios de crecimiento profesional y portafolios de proceso, entre otros².

Así, se puede encontrar una amplia tipología de portafolios. Meeus y Lois (2006) distinguen hasta veintiocho clasificaciones de portafolios. Algunas de ellas se hacen en función del autor: discente o docente; del registro: papel o digital; de la defensa: escrito u oral; y del propósito: para la certificación, para la selección o para la promoción. Otra clasificación se hace en función del contenido, del momento de desarrollo del portafolios y de los objetivos que se persiguen (Barragán, 2005, citado por Pérez, Alcaraz y Fernández, 2016; Starkie, 2009).

De ahí que Barberá (2005) exponga que: “en la bibliografía encontraremos tantas definiciones y desarrollos de portafolios como autores seamos capaces de consultar (p. 499, citado por Pérez, Alcaraz y Fernández, 2016).

En este trabajo trataremos de resumir las aportaciones revisadas en dos categorías generales teniendo en cuenta que no se trata de una escala de negro o blanco, sino que también podemos encontrar matices.

La primera categoría, el primer tipo de portafolios que destacaremos, será el portafolios de producto (traducido como *showcase*)³. La segunda categoría o tipo sería el portafolios de proceso.

En el de producto, *showcase portfolio*, deja de prevalecer, y así lo expone Pollari (2000), el proceso. Lo que importa es el resultado final que se ha de exponer a una determinada audiencia. En el otro, en el de proceso, cobra protagonismo todo el proceso de enseñanza y aprendizaje, que trata de recogerlo pero que debe ir acompañado siempre de la comprensión, la reflexión conjunta y la autoevaluación del sujeto que aprende. Sin embargo, tanto uno como otro suelen usarse para obtener un mérito final, lo cual nos lleva a diferenciar los tipos pero a mezclar las funciones.

2.3. Algunos usos del portafolios

Pérez, Alcaraz y Fernández (2016) recogen la lista que Klenowski (2004) expone con los diferentes usos que se pueden encontrar del portafolios en el ámbito educativo:

² Además de estas clasificaciones, hay que mencionar dos categorías más de portafolios que dependen del soporte: papel o virtual. El uso del e-portafolios ha proliferado mucho en la última década. Para profundizar en este, puede acudir a Hernández (2008), Barret (2009) y Soto (2016), entre otros.

³ Término que los anglosajones utilizan para referirse al portafolios de producto.

- ✓ El portafolios para la valoración del rendimiento, lo que denomina evaluación sumativa. Debe entenderse como calificación. Este acumula el material recogido durante un periodo de tiempo determinado.
- ✓ El portafolios para conseguir una certificación o para la selección de profesionales es aquel que contiene una selección de los trabajos más representativos del aspirante.
- ✓ El portafolios que permite promocionar dentro de un ámbito profesional. En función de los criterios de valoración expuestos por el contratante, el sujeto que aspira a ser promocionado selecciona los trabajos más acordes a tales criterios.
- ✓ El portafolios para el desarrollo profesional. Se refiere al portafolios de los-los docentes, cuyo fin es reflexionar críticamente sobre la propia práctica.
- ✓ Por último, el portafolios para la evaluación y el aprendizaje, cuyo objetivo es ayudar al-la estudiante a desarrollar y profundizar sobre su proceso de aprendizaje.

Si profundizamos en las finalidades de cada uno, descubriremos que, de todos los posibles usos enumerados, solo uno de ellos no está asociado a la rendición de cuentas. Nos referimos al portafolios para el desarrollo profesional o, lo que es lo mismo, el portafolios del-la docente. El resto, incluido el portafolios para la evaluación y el aprendizaje, tiene un objetivo final que es demostrar algo. Con lo cual podríamos hablar de dos usos fundamentales: para rendir cuentas y para el desarrollo personal o profesional, en definitiva, para el aprendizaje.

3. Condiciones de un portafolios para la evaluación y el aprendizaje

Dada la relativa novedad del uso de portafolios en el ámbito educativo –a la que venimos haciendo referencia a lo largo del artículo–, nos encontramos ante el reto de tratar de dotarlo de sentido educativo, situándolo al servicio de la evaluación. Es decir, desplazar la función sumativa del mismo y resaltar la formativa.

Veremos a continuación las características o condiciones que un portafolios educativo debe tener o cumplir para convertirlo en esa “ocasión de aprendizaje” a la que se refería Álvarez Méndez (2008). O lo que es lo mismo, poner a la evaluación al servicio del aprendizaje para lo cual es condición necesaria, tal y como expone Alcaraz (2014), que esta fuese respetuosa con los principios psicológicos y pedagógicos que sustentan la visión constructivista del currículum y del aprendizaje en el aula, alejándose de la concepción tecnocrática de la enseñanza y suponiendo fuente de mejora para quien aprende y para quien enseña.

3.1. El portafolios debe centrarse en el análisis y reflexión de los procesos y contextos de aprendizaje

Tal y como defienden Román y Murillo (2009); Pérez Gómez, Soto, Sola y Serván (2009); Álvarez Méndez (2001); Alcaraz, Fernández y Sola (2012); Murillo e Hidalgo (2015), entre otros, la evaluación ha de servir para la mejora del aprendizaje del alumnado, es decir, la evaluación es también aprender cómo aprender. Pérez Gómez, Soto, Sola y Serván (2009) exponen:

De requisito de control burocrático, mecánico y externo que sanciona los rendimientos, establece calificaciones y clasifica a los estudiantes se convierte en un proceso complejo, cualitativo, de conocimiento de los propios procesos de aprendizaje, reflexión sobre ellos

planificación individual y compartida de proyectos de mejora, donde al final la evaluación incorporada como una cultura se convierte en el mejor instrumento de aprendizaje mediante la reflexión en la acción y sobre la acción. (209, p. 6)

Es a lo que Murillo e Hidalgo (2015) se refieren cuando expresan que una evaluación para la justicia social es aquella que, entre otras cuestiones, pasa de ser una evaluación “del” aprendizaje a una evaluación “para” el aprendizaje. De modo que desde esa perspectiva el portafolios debe servir “para” el aprendizaje; para auto-gestionar el propio proceso de aprendizaje, puesto que quien aprende evalúa a su vez, discrimina, valora, critica, opina, razona, fundamenta y decide entre lo que tiene valor y lo que no para el proceso de aprender. Por ello, el-la docente que pone en práctica la evaluación a través de portafolios debe preocuparse por incluir procesos de organización, de autoevaluación, de metacognición y de reflexión. Hebert (2001), Barberá (2005), Klenowsky (2004), Barret (2001), Lyons (1999) y Soto (2016), consideran a este último el elemento clave en el trabajo con portafolios.

De acuerdo con Soto (2016), la reflexión no se da de forma espontánea, sino que necesita aprenderse. Para Soto (2016) y Le Boterf (2000), la reflexión implica que la persona se distancie de su actividad para poder contemplar el problema y sus variables tratando de descontaminarse de su inmediatez. A ello es a lo que denominan metacognición.

La reflexión sobre los procesos y contextos en los que tiene lugar el aprendizaje, no solo del-la alumno-a sino del-la maestro-a como una condición necesaria a la que se debe aspirar con el desarrollo de un portafolios, convierte a la evaluación en una oportunidad de aprendizaje para el sujeto que aprende y para el que enseña. De modo que tomar consciencia, conocerse a sí mismo y saber cómo uno aprende deben ser aspiraciones de la práctica del portafolios. Es decir, es necesario estimular la metacognición; de acuerdo con Pérez Gómez (2012, p. 210), “la metacognición como medio para promover la capacidad de autonomía y autorregulación del aprendizaje y del desarrollo”. Wenger (2010), citado por Pérez Gómez (2012), propone considerar tres aspectos fundamentales del diseño como autorregulación que cobran mucho sentido en el desarrollo de un portafolios educativo: una buena orientación, una buena reflexión y una buena exploración de las nuevas posibilidades.

También la autoevaluación, según Pérez Gómez, Soto Gómez, Sola Fernández y Serván (2009), es un motor potente para desarrollar la metacognición: la autoevaluación como mecanismo de aprendizaje que implica organizarse, planificarse, reflexionar, etcétera.

En este mismo sentido, la organización del aprendizaje y para el aprendizaje es otra de las claves. Ordenar el trabajo que se va haciendo contribuye a organizar el pensamiento, a reconocer lo que se lleva trabajado y lo que queda por trabajar.

La organización cobra protagonismo en la medida en que el portafolios debe tratar de presentar, representar el proceso de trabajo, la construcción del conocimiento, las idas y venidas del-la estudiante.

En definitiva, el portafolios asume la condición de estar atento al proceso y al contexto en el que se produce el aprendizaje, en tanto en cuanto debe favorecer la reflexión, la autoevaluación, la metacognición, la organización y la autorregulación del conocimiento que se construye sobre el mundo y sobre uno-a mismo-a. En este sentido, podemos definir al portafolios como la trastienda del aprendizaje.

3.2. El portafolios debe fomentar un clima de libertad que propicie la comunicación y el aprendizaje auténticos

Que el-la estudiante asuma responsabilidades debe ser cualidad de una buena práctica con portafolios. Para ello se necesita un clima de libertad donde la autonomía de quien está aprendiendo vaya creciendo gradualmente. Autonomía del-la alumno-a combinada con guía del-la docente. El-la estudiante debe poder decidir, lo cual pasa por crear parcelas de libertad, de negociación, de argumentación y diálogo, alejadas de intenciones que pretendan oprimir, castigar, silenciar o humillar, más propias de la calificación que de ninguna otra práctica de la escuela.

El portafolios que aspira a evaluar debe impregnar el aula de confianza, liberándola de amenazas, sustituyendo el poder que otorga la calificación por la autoridad de la que nos dota la evaluación. Es así como lograremos poco a poco que tengan lugar una comunicación y un aprendizaje auténticos.

Al referirnos a la condición de autenticidad de la comunicación, hablamos de que el portafolios como práctica de evaluación debe dirigir todos sus esfuerzos a que la información que nos da el alumnado sea lo más sincera posible. Asimismo, el portafolios que sirve al aprendizaje implica la comunicación fluida, relevante y veraz entre docente y estudiante. Cobrando especial importancia el papel del-la docente que debe presentarse a los-as alumnos-as como un guía permanente que propicie confianza y seguridad (Gallardo, Sierra, Barquín y Medeiros, 2016). Tal y como acentúan Calero y Calero (2008), el uso del portafolios desde una perspectiva vigostkiana del aprendizaje refuerza el papel del-la maestro-a como figura mediadora, guía, modeladora y facilitadora del aprendizaje y del contexto que lo fomente.

Del mismo modo, el desplazamiento de la calificación, la no constante presencia de la misma en la vida del aula, contribuye a recuperar el tan maltrecho valor de uso del conocimiento que en tantas ocasiones ha sido relegado por el valor de cambio. Trabajar con portafolios en las aulas no debería implicar la presencia continua de pequeñas notas que van obteniendo cada una de las actuaciones del-la estudiante, que fragmentan el aprendizaje en pequeños “logros”. Al contrario, debe ser una herramienta de aprendizaje a largo plazo, que ponga al alumnado en situación de construir y reconstruir su trayectoria en el proceso de aprender, con la mediación del-a profesor-a, y respetuosa con los ritmos de maduración de cada alumno-a. Se aboga así por un aprendizaje más auténtico, no tan mediatizado por las recompensas de la calificación.

Por tanto, el portafolios no debe ser –como no debe serlo el conocimiento– una mera mercancía sin ninguna utilidad más allá de recopilar los “logros” para ser expuestos y valorados por el profesorado y, de este modo, intercambiados por una calificación.

3.3. El portafolios debe apostar por el error como fuente de aprendizaje y desplazar los juicios penalizadores

Una de las condiciones de la evaluación es la de preocuparse por los errores del alumnado, no en el sentido de penalizarlos, sino en el sentido de comprenderlos y transformarlos. Como exponen Murillo e Hidalgo (2015), en lugar de ser un mecanismo de control y castigo, la evaluación debe centrarse en reforzar lo positivo y en servir de ayuda al alumnado. Debemos practicar, en palabras de estos autores, una “evaluación optimista”. También Postman (2001) reconoce que la historia del aprendizaje es la narración de nuestros errores superados. Por tanto, el portafolios debe servir para buscar las respuestas a los problemas que el alumnado se va encontrando a

lo largo de su proceso de aprendizaje, así como para solucionar los errores cometidos y sobreponerse a las adversidades.

El portafolios debe ayudar al alumnado a construir una mente informada y organizada de lo que aprende y cómo lo aprende. En este sentido, cobra relevancia la corrección informada, es decir, comentarios que orienten a quien está en situación de aprendizaje, ayudándole a tomar consciencia de sus puntos fuertes y sus puntos débiles y de cómo fortalecer estos últimos.

En definitiva, la corrección informada es, en gran parte, la antesala a la reflexión de la que hablábamos anteriormente.

3.4. El portafolios debe ser un canal que recoja la voz del alumnado y ayude a la comprensión holística de la vida del aula

El portafolios, para que sea útil a la evaluación, debe permanecer atento a la voz y al sentir del alumnado. Es sin duda el-la alumno-a quien construye su aprendizaje y por tanto quien va dando forma y sentido al proceso de creación del portafolios. En este sentido, está bastante claro que el portafolios tiene que dar voz al-la estudiante, porque es “su” voz la que debe sonar. Pero reconocer la voz y la participación del alumnado en el aula no es solo escuchar lo que tiene que decir sobre sí mismo-a, sobre su propio proceso; es animarle a que exprese también cuestiones sobre el contexto que rodea a ese proceso.

Los-as alumnos-as tienen opiniones sobre las condiciones en las que tiene lugar la enseñanza y el aprendizaje; tienen opiniones de los-as docentes, de los materiales, de las actividades, de los tiempos, de los agrupamientos, del clima de clase, entre otras. Unas veces más sustanciosas y otras menos. Ahí es donde juega un papel fundamental el portafolios, que debe fomentar la participación activa en el aula. El-la maestro-a que se atreve con un portafolios debe atreverse con la democracia. Preguntar, escuchar, debatir, dialogar, argumentar, negociar cualquier aspecto relacionado con la vida escolar, es también trabajar con portafolios. Este debe ser un ejercicio de alimentar la palabra escrita pero también la palabra hablada. El portafolios al servicio de la evaluación es un ejercicio democrático.

Para ello, los-as docentes deben abrir vías de comunicación constantes que recojan manifestaciones orales, escritas, anónimas, nominales, individuales y grupales sobre los variados elementos que confluyen en el día a día del aula, no solo dando la palabra al-la alumno-a, sino ayudándole a expresarla y a utilizarla; lo cual requiere dedicación, tiempo y paciencia.

3.5. El portafolios debe convertirse en un sistema de evaluación

Llegados a este punto, tras haber expuesto todo lo que el portafolios debe ser a la vez, se hace necesario resaltar la necesidad de convertir a este en un sistema de evaluación compuesto por diferentes herramientas que conectadas entre sí deben transformar la práctica educativa del aula.

El portafolios no es un medio físico a través del cual el alumnado presenta su trabajo y el profesorado lo valora. Para que realmente este sirva a la evaluación y por tanto al aprendizaje, debe ser una manera de entender la práctica educativa, una forma de vida en el aula, un sistema complejo que pretenda abarcar todos los elementos influyentes en el proceso de enseñanza-aprendizaje y que trate de arrojar luz sobre todos ellos.

Pérez Gómez y Serván (2016) defienden la idea del portafolios no como instrumento de evaluación sino como sistema que puede contener otros instrumentos y recursos que han de coordinarse.

Para que realmente el portafolios tenga sentido educativo no debería ser reducido a la categoría de instrumento, sino abrirse, complejizarse en el sentido de incluir un conjunto de herramientas y estrategias puestas en práctica con una finalidad: la de comprender y mejorar el proceso de enseñanza-aprendizaje. Para ello, es necesario incorporar una serie de elementos que nos permitan abarcar y analizar de forma holística la realidad del aula. Alcaraz, Mayorga y Pérez (2016) exponen la importancia de incluir en los portafolios educativos material de anclaje elaborado por el-la docente, como tablas, fichas, cuestionarios, que sirvan de guía al alumnado para preguntarse y reflexionar sobre el proceso en el que se encuentra, sobre cómo ha llegado hasta ahí y cómo ha de representarlo. Veremos a continuación algunas sugerencias de herramientas y estrategias que pueden integrar la práctica de evaluar a través de portafolios:

- ✓ *Diarios*: A través de los diarios el profesorado puede ir promoviendo gradualmente el nivel de reflexión del alumnado que suele estar poco habituado a ello. A medida que los-as estudiantes trabajan en solitario o en grupo, construir un diario individual o colectivo puede resultar una oportunidad para que estos-as vayan incorporando poco a poco elementos de análisis sobre lo que aprenden y cómo lo aprenden. Es importante aquí no desesperar si la calidad de las primeras reflexiones es baja y que el-la docente se configure como un elemento de apoyo en ese camino para que poco a poco el nivel de profundidad y argumentación vaya creciendo.
- ✓ *Cuestionarios de autoevaluación*: Plantear al alumnado preguntas sobre cuestiones relativas a su propio proceso de aprendizaje puede resultar útil para que poco a poco este vaya tomando consciencia de lo que ha trabajado, de lo que le ha resultado más o menos difícil, de lo que más le ha gustado, de lo que le queda por mejorar, de los compromisos que adquiere, etcétera. Incorpora la valoración cualitativa que el-la alumno-a hace de su trabajo. Las preguntas pueden empezar siendo sencillas y, a medida que se va avanzando, estas pueden requerir respuestas algo más elaboradas. El cuestionario de autoevaluación es un recurso útil para el sujeto que aprende y para quien enseña, puesto que de este modo el-la profesor-a “accede” a lo que el-la alumno-a piensa sobre su proceso de aprendizaje. Para que este tenga sentido es necesario demostrar al alumnado que no se trata de auto-ponerse nota y que ello no tiene implicaciones en la calificación final ya que, si no, es muy probable que este responda en función de lo que cree que queremos oír. Es recomendable que no sea anónimo para que de este modo el-la docente pueda detectar necesidades.
- ✓ *Fichas u hojas de registro*: Se trata de poner al alumnado en situaciones en las que deba ir anotando el trabajo que se lleva hecho. Es el lugar (puede ser una tabla compuesta por filas y columnas) donde se registran los proyectos realizados, las actividades desarrolladas, etcétera, de forma sintética. Para que tenga sentido educativo, puede ir acompañada de observaciones tanto del-la estudiante como del-la profesor; de este modo, se da lugar a una especie de diálogo entre las partes ya que el-la alumno-a registra un comentario acerca de un proyecto o actividad y el-la docente puede contestarle por escrito.

- ✓ *Cuestionarios de evaluación:* Se trata de un listado de preguntas que el-la docente hace al alumnado sobre aspectos generales o concretos del desarrollo de clase. Debe contestarse de forma anónima dado que el sentido de este es conocer el punto de vista del alumnado, la opinión que este construye acerca del modo en que se trabaja, del clima de clase o de la actitud del-a profesor-a. Para ello, es importante crear un clima de confianza y seguridad en el que los-as estudiantes puedan expresarse con libertad. Se trata de incorporar la voz del alumnado sobre lo que acontece en el aula, de escucharle e incorporar a nuestro quehacer diario el modo en que los protagonistas del mismo perciben y entienden nuestras actuaciones.
- ✓ *Buzones de Freinet:* Son los conocidos buzones “propongo”, “felicitó” y “crítico”, cuya finalidad es animar al alumnado a dejar mensajes a compañeros-as o maestros-as del aula, incluso del centro. Felicitar a alguien, criticar una acción o proponer algo es otro modo de formar parte activa de la vida escolar. El mecanismo es sencillo: los buzones deben permanecer fijos en el aula y disponibles para ser usados en cualquier momento. Se debe planificar un día (puede ser uno a la semana o cada quince días) para la lectura de los mensajes. Como todas las anteriores, se trata de una práctica que hay que trabajar y alimentar dentro de un clima de seguridad, confianza y afecto.
- ✓ *Corrección de las producciones del alumnado:* Se trata de incorporar al portafolios las producciones del alumnado, con objeto de ser leídas y corregidas por el-la docente que debe proporcionar retroalimentación a sus autores-as, de modo que estos-as puedan reconducir el trabajo, incorporar nuevas cuestiones, ampliar o reducir, atendiendo a las sugerencias del-la profesor-a.

En definitiva, destacamos que tal y como afirman Calero y Calero (2008):

Entendemos que el valor real del portafolios que planteamos no reside en su apariencia física o en su organización concreta, sino en el contexto que crea de desarrollo y de cooperación que generan los procesos de reflexión y colaboración entre profesor y alumno.
(p.32)

4. Concluyendo

A lo largo de este artículo se han expuesto no solo el origen de las prácticas con portafolios en el ámbito educativo como herramienta de evaluación, sino algunos de los diferentes tipos y usos que se desprenden de dicha práctica. A su vez, en este trabajo se han descrito una serie de condiciones que un portafolios al servicio de la evaluación debería cumplir. Entendiendo la evaluación como proceso de recogida de información, planificado y horizontal, que no solo recoge sino que devuelve información que sea útil para conocer, comprender y analizar lo que influye en el proceso de enseñanza-aprendizaje, con la intención de mejorarlo (Alcaraz, 2015b).

También se ha resaltado la relativa novedad del uso de portafolios en el ámbito educativo, especialmente en la etapa obligatoria, por lo que es importante destacar que, ahora más que nunca, nos encontramos ante el reto de tratar de dotarlo de sentido educativo, situándolo al servicio de la evaluación. Pero ¿qué significa esto? Significa que, como hemos ido exponiendo, debemos convertirlo en un medio y en un fin para el aprendizaje. En una oportunidad para analizar y reflexionar sobre los procesos y los contextos de aprendizaje; en un medio para propiciar un clima de libertad, participación y comunicación auténtica. En una vía donde el error sea considerado fuente de

aprendizaje. Es decir, en un sistema, compuesto a su vez por otras herramientas o instrumentos que obedezcan a una filosofía concreta de entender y vivir la evaluación en el aula. Descartando la función sumativa y resaltando la formativa, ya que podemos correr el riesgo de trasladar el uso de rendición de cuentas que del portafolios se ha empezado a hacer en contextos de educación superior, donde la función selectiva del sistema es mucho más clara y acentuada que en la enseñanza obligatoria.

De sobra es conocido que no son las herramientas en sí mismas las que limitan o posibilitan mejores o peores condiciones de aprendizaje, sino que son sus usos. Usos que dependen directamente de las concepciones de aprendizaje, enseñanza, currículum y evaluación de la persona que las pone en práctica. De modo que cualquier herramienta puede estar al servicio de la calificación o al de la evaluación. Así, por ejemplo, el examen podría estar al servicio de una u otra según cómo se use y para qué se use. Sin embargo, y aunque es probablemente cierto que lo que conocemos por examen pueda emplearse de una forma u otra en función de quién lo emplea, ocurre particularmente con el examen que, al ser una herramienta con tan larga trayectoria calificativa, la línea entre utilizarlo para calificar y evaluar es sumamente fina. Por este motivo, insistimos en la crítica hacia los ‘usos’ y ‘desusos’ del examen como herramienta de calificación y/o evaluación ya que, si queremos contribuir a ofrecer claridad al debate evaluación-calificación, apostar por la evaluación sin desprendernos de los exámenes –pesos pesados de la calificación–, es arriesgarnos a seguir mezclando todo.

Por eso, insistimos en que una de las ventajas del portafolios es precisamente su novedad en el ámbito de la pedagogía; cualidad que no lo exime de posibles riesgos, tal y como apuntábamos antes (la trayectoria en otros contextos como repositorio que “demuestra” nuestros méritos o capacidades). Por lo tanto, es importante que la cualidad continua del desarrollo de un portafolios para evaluar en el aula no se convierta en la barita mágica con la que tratemos de “medir” constantemente las actitudes, acciones, manifestaciones, etcétera, del alumnado. Como plantea Pérez Gómez (2012, p. 96): “Las recomendaciones de la evaluación continua se han malentendido o convertido en exigencias de examen o calificación continua”.

En definitiva, podemos decir que entre la ventaja y el riesgo asumimos un reto: aprovechar el carácter continuo y reflexivo que se le otorga en la literatura al portafolios para convertirlo en una oportunidad de evaluación “para” el aprendizaje, más que “sobre” el aprendizaje, ya que caer en la tentación de usar el portafolios para puntuar constantemente lo que los-as estudiantes realizan en clase sería desperdiciar una buena ocasión para plantear en el aula una auténtica evaluación .

Referencias

- Alcaraz, N. (2014). Un viejo trío de conceptos: Aprendizaje, currículum y evaluación. *Aula De Encuentro*, 2(16), 55-86.
- Alcaraz, N. (2015a). Aproximación histórica a la evaluación educativa: De la generación de la medición a la generación ecléctica. *Revista Iberoamericana de Evaluación Educativa*, 8(1), 11-25.
- Alcaraz, N. (2015b). Evaluación vs. calificación. *Aula de Encuentro*, 17(2), 209-236.

- Alcaraz, N., Fernández, M. y Sola, M. (2012). La voz del alumnado en los procesos de evaluación docente universitaria. *Revista Iberoamericana de Evaluación Educativa*, 5(2), 27-39.
- Álvarez Méndez, J. M. (2001). *Evaluar para conocer, examinar para excluir*. Madrid: Morata.
- Álvarez Méndez, J. M. (2008). Evaluar el aprendizaje en una enseñanza centrada en las competencias. En J. Gimeno Sacristán (Comp.), *Educación por competencias, ¿qué hay de nuevo?* (pp. 206-234). Madrid: Morata.
- Barberá, E. (2005). La evaluación de competencias complejas: La práctica del portafolio. *Educere*, 9(31), 497-503.
- Barragán, R. (2005). El portafolio, metodología de evaluación y aprendizaje de cara al nuevo Espacio Europeo de Educación Superior. Una experiencia práctica en la Universidad de Sevilla. *Revista Latinoamericana de Tecnología Educativa*, 4(1), 121-139.
- Barret, H. (2001). Electronic portfolios. Educational technology. En *Educational Technology Encyclopedia*, ABC-CLIO. Recuperado de <http://electronicportfolios.com/portfolios/encyclopediaentry.htm>
- Barret, H. (2010). Balancing the two faces of eportfolios. *Educação, Formação y Tecnologías*, 3(1), 6-14.
- Bastidas, J. A. (1996). The teaching portfolio. A tool to become a reflective teacher. *English Teaching Forum*, 34(3-4), 24-48.
- Calero, A. y Calero, E. (2008). El portafolio como estrategia de evaluación compartida de la comprensión lectora. *Didáctica. Lengua y Literatura*, 20, 15-36.
- Calfee, R. C. y Freedman, S. W. (1996). Classroom writing portfolios: Old, new, borrowed, blue. En R. Calfee y P. Perfumo (Eds.), *Writing portfolios in the classroom* (pp. 3-26). Mahwah, NJ: Lawrence Erlbaum Associates.
- Danyelson, C. y Abrutyn, L. (1999). *Una introducción al uso del portafolios en el aula*. México D. F.: Fondo de Cultura Económica.
- Darling-Hammond, L. (2005). Standards and assessments: Where we are and what we need. *Teachers College Record. The Voice of Scholarship in Education*. Recuperado de <http://www.tcrecord.org/PrintContent.asp?ContentID=11109>
- Delmastro, A. L. (2005). El uso del portafolios en la enseñanza de las lenguas extranjeras. Perspectiva del docente. *Investigación y Postgrado*, 20(2), 187-211.
- Díaz Barriga, F. (2006). *Enseñanza situada: Vínculo entre la escuela y la vida*. México D. F.: McGraw-Hill.
- Díaz Barriga, F. y Pérez, M. (2010). El portafolio docente a escrutinio: Sus posibilidades y restricciones en la formación y evaluación del profesorado. *Observar*, 4, 6-27.
- French, R. L. (1992). Portfolio assessments and LEP students. En T. D'Emilio y S. Coleman (Coord.), *Focus on evaluation and measurement: Proceedings of the Second National Research Symposium on Limited English Proficient Student Issues* (pp. 249-285). Washington D. C.: United States Department of Education, OBEMLA.
- Gallardo, M., Sierra, E., Barquín, F. J. y Medeiros, M. (2016). Despertando las miradas: Propósitos, criterios y sentido del portafolios educativo. En A. I. Pérez Gómez (Dir.), *El portafolios educativo en Educación Superior* (pp. 65-78). Madrid: Akal.
- García Hernández, E. (2000). *Algunas aplicaciones del portafolios en el ámbito educativo*. Chihuahua: Secretaría de Educación y Cultura del Estado de Chihuahua. Recuperado de http://sgpwe.izt.uam.mx/files/users/virtuami/file/ext/gestion_evaluacion_actv_ext_portafolios.pdf

- Hamilton, S. J. (1994). Portfolio pedagogy: Is a theoretical construct good enough? En L. Black, D. A. Daiker, J. Sommers y G. Stygall (Eds.), *New directions in portfolio assessment* (pp. 157-167). Portsmouth, NH: Boynton/Cook Publishers, Heinemann.
- Hebert, E. (2001). *The power of portfolios. What children can teach us about learning and assessment*. San Francisco, CA: Jossey-Bass.
- Hernández Hernández, F. (2008). El portafolios electrónico y el aprendizaje para la comprensión en Secundaria. *Investigación en la Escuela*, 64, 59-71.
- Hidalgo, J. y De la Blanca, S. (2013). La contribución del prácticum a la innovación metodológica en la educación de adultos. En P. C. Muñoz Carril, M. Raposo Rivas, M. González Sanmamed, M. E. Martínez Figueira, M. A. Zabalza Cerdeiriña y A. Pérez Abellás (2013), *Un prácticum para la formación integral de los estudiantes* (pp. 943-953). Santiago de Compostela: Andavira.
- Jalbert, P. (1998). *Le portfolio au service de l'apprentissage et de l'évaluation*. Montréal-Toronto: Chenelière/McGraw-Hill.
- Klenowski, V. (2004). *Desarrollo del portafolios para el aprendizaje y la evaluación. Procesos y principios*. Madrid: Narcea.
- Le Boterf, G. (2000). *Construire les compétences individuelles et collectives*. París: Les Editions d'Organisation.
- Lyons, N. (Comp.). (1999). *El uso de portafolios. Propuestas para un nuevo profesionalismo docente*. Buenos Aires: Amorrortu.
- Meeus, P. y Lois, V. (2006). Portfolio in higher education: Time for clarifactory framework. *International Journal of Teaching and Learning in Higher Education*, 17(2), 127-135.
- Monereo, C. (2009). La autenticidad de la evaluación. En M. Castelló (Coord.), *La evaluación auténtica en Enseñanza Secundaria y Universitaria* (pp. 9-22). Barcelona: Edebé.
- Murillo, F. J. e Hidalgo, N. (2015). Dime cómo evalúas y te diré qué sociedad construyes. *Revista Iberoamericana de Evaluación Educativa*, 8(1), 5-9.
- Murphy, S. y Camp, R. (1996). Moving toward systemic coherence: A discussion of conflicting perspectives on portfolio assessment. En R. Calfee y P. Perfumo (Eds.), *Writing portfolios in the classroom* (pp. 103-148). Mahwah, NJ: Lawrence Erlbaum Associates.
- Oosterhof, A. (1996). *Classroom applications of educational measurement*. Columbus, OH: Merrill.
- Paulson, F., Paulson, P. y Meyer, C. (1991). What makes a portfolio a portfolio. *Educational Leadership*, 60-63.
- Pelavin, S. (1991). *Performance assessments in the states*. Washington, DC: Pelavin Associates.
- Pérez Gómez, A. I. (2012). *Educarse en la era digital*. Madrid: Morata.
- Pérez Gómez, A. I., Soto Gómez, E., Sola Fernández, M. y Serván, M. J. (2009). *La evaluación como aprendizaje. ¿Evaluación = Calificación?* Madrid: Akal.
- Pérez Gómez, A. I. y Serván, M. J. (2016). Naturaleza y sentido del portafolios educativo. En A. I. Pérez Gómez (Dir.), *El portafolios educativo en Educación Superior* (pp. 39-45). Madrid: Akal.
- Pérez, L., Alcaraz, N. y Fernández, M. (2016). Revisión bibliográfica sobre el portafolios. En A. I. Pérez Gómez (Dir.), *El portafolios educativo en Educación Superior* (pp. 17-38). Madrid: Akal.
- Pérez Rendón, M. (2014). Evaluación de competencias mediante portafolios. *Perspectiva Educativa*, 53(1), 19-35. doi:10.4151/07189729-Vol.53-Iss.1-Art.213

- Pollari, P. (2000). *This is my portfolio: Portfolios in upper secondary school English studies*. Recuperado de <http://eric.ed.gov/PDFS/ED450415.pdf>
- Postman, N. (2001). *El fin de la educación. Una nueva definición del valor de la escuela*. Barcelona: Octaedro.
- Reckase, M. (1995). Portfolio assessment: A theoretical estimate of score reliability. *Educational Measurement: Issues and Practice*, 14(1), 12-14. doi:10.1111/j.1745-3992.1995.tb00846.x
- Román, M. y Murillo, F. J. (2009). La evaluación de los aprendizajes escolares. Un recurso estratégico para mejorar la calidad educativa. *Revista Iberoamericana de Evaluación Educativa*, 1(2), 5-9.
- Shores, E. F. y Grace, C. (2004). *El portafolios paso a paso: Infantil y primaria*. Barcelona: Graó.
- Soto, E. (2016). La reflexión corazón y alma del portafolios educativo. En A. I. Pérez Gómez (Dir.), *El portafolios educativo en Educación Superior* (pp. 97-118). Madrid: Akal.
- Starkie, E. (2009). El papel del docente y del alumno en la elaboración de portafolios. *Revista de Educación*, 349, 451-462.
- Villalobos, J. (2002). Portafolios y reflexión: Instrumentos de evaluación en una clase de escritura. *Educere*, 5(16), 390-396.
- Wenger, E., White, N. y Smith, J. (2010). *Digital habitats, stewarding technology for communities*. Portland, OR: CPsquare.
- Wiggings, G. (1990). *The case for authentic assessment*. Recuperado de <http://www.ericdigests.org/pre-9218/case.htm>
- Wile, J. M. y Tierney, R. J. (1996). Tensions in assessment: The battle over portfolios, curriculum and control. En R. Calfee y P. Perfumo (Eds.), *Writing portfolios in the classroom* (pp. 203-215). Mahwah, NJ: Lawrence Erlbaum Associates.

Análisis de una Escala para Evaluar la Satisfacción del Profesional Docente en la Enseñanza Primaria

Analysis of a Scale to Evaluate the Satisfaction of the Educational Professional in the Elementary Education

Josefa Rodríguez Pulido, Jaime León González-Vélez,
María Victoria Aguiar Perera* y Josué Artilés Rodríguez

Universidad de Las Palmas de Gran Canaria

Este trabajo forma parte de un estudio promovido por el grupo de investigación y de innovación pedagógica de la Universidad de Barcelona y el Observatorio Internacional de la Educación de la Profesión Docente. Elaborar una escala de valoración de la satisfacción docente (ESD) es el objetivo de esta investigación, por ello se determina si el instrumento desarrollado y las inferencias realizadas en este cuestionario son válidas. La escala se aplica a una muestra de 215 profesores de primaria en la Comunidad Autónoma de Canarias (Gran Canaria). El cuestionario propuesto por el grupo de FODIP, consta de 18 ítems tipo Likert de 1 a 5. Los resultados muestran que la escala de valoración aportada es válida y fiable, y que su estructura de tres factores (la planificación, la autonomía y el desarrollo profesional) es una alternativa ágil y rápida para conocer el grado de satisfacción del profesorado. Las tres dimensiones de la escala encuentran correlatos similares en modelos anteriores y, aunque como ya señalamos, el número de dimensiones estimado por los distintos investigadores es variado, a juzgar por las referencias, una escala corta y de fácil aplicación puede ayudar a una primera aproximación a la satisfacción docente.

Palabras clave: Satisfacción del profesional docente, Enseñanza primaria, Escala, Planificación, Autonomía, Desarrollo profesional.

This work is part of a study promoted by the Group of research and teaching innovation of the University of Barcelona and the international Observatory of the education of the profession teaching. To elaborate a scale of valuation of the educational satisfaction (ESD) is the aim of this investigation, for it decides if the developed instrument and the inferences that realized in this questionnaire are valid. The scale is applied to a sample of 215 primary teachers in the autonomous community of Canary Islands (Gran Canaria). The questionnaire to assess the satisfaction of professional teachers in primary education proposed by the Group of FODIP, it consists of 18 items Likert-type 1 to 5. The results show that the scale of valuation provided is valid and reliable, and that its structure of three factors (planning, autonomy and professional development) is an agile and fast alternative to knowing the degree of satisfaction of the teaching staff. The three dimensions of the scale are similar correlates in previous models and, although as already noted, the number of dimensions estimated by different researchers is varied, judging by the references, a scale short and easy application can help a first approach to teacher satisfaction.

Keywords: Satisfaction of the educational professional, Elementary education, Scale, Planning, Autonomy, Professional development.

1. Introducción

Elaborar una escala de valoración de la satisfacción docente (ESD) es la aportación de esta investigación. Nuestra contribución viene apoyada por la satisfacción en el trabajo educativo como una dimensión importante de la salud de la organización y desarrollo de los estudiantes. La indagación “del profesor ideal ha sido una constante a lo largo de la historia, su estilo de enseñanza es un valor importante en todo proceso educativo dada la repercusión de la organización del material de estudio, su forma de comunicarse, la directividad de la clase, su rol orientador, consejero, facilitador del aprendizaje y transformador de la cultura, su nivel de control de la clase, su imparcialidad, empatía, entusiasmo y su disponibilidad afectiva. Todas ellas son cualidades personales consideradas como prerequisites para el éxito de la enseñanza. No podemos olvidarnos de que el fenómeno educativo, a pesar de sus implicaciones sociales, es un fenómeno individual” Peiteado (2013, p.66). De ahí que la emoción positiva que el profesor tenga en relación con la evaluación de su actividad docente nos podría dar evidencias, del presente y del momento, de su nivel de satisfacción (De Pablos, González y González, 2008).

Estudios realizados por Cuscó y Wells (2009) ponen de manifiesto que la insatisfacción del profesorado, viene dada por la realización de muchas tareas ajenas a la profesión y por no tener condiciones para enseñar.

Este trabajo forma parte de un estudio promovido por el grupo de investigación y de innovación pedagógica de la Universidad de Barcelona y el Observatorio Internacional de la Educación de la Profesión Docente (FODIP). El uso de escalas fiables y válidas es esencial para evaluar la satisfacción del trabajo del profesor. Las investigaciones de satisfacción docente han sido planteadas desde diferentes posicionamientos, se realizan investigaciones sobre la satisfacción laboral del docente, del alumnado, del director e incluso en el marco de la educación infantil, primaria, secundaria. Ello conlleva que la terminología utilizada sea variada (malestar docente, salud mental del profesorado, burnout docente, angustia del enseñante o estrés y ansiedad del profesorado).

Las investigaciones realizadas han profundizado no sólo para mejorar la satisfacción personal, sino también para mejorar los resultados de la organización. Satisfacción es un elemento prioritario para evaluar y mejorar la calidad de los maestros, el proceso de formación y el desarrollo organizativo en general (Anaya y Suárez, 2007). Por otra parte, la investigación sobre la satisfacción también es importante para el desgaste y la separación de los docentes (Skaalvik y Skaalvik, 2011).

Desde los primeros estudios desarrollados por Hoppock (1935), se ha ido avanzando hacia estudios relacionados con la interacción de variables personales y ambientales (Carr, 2007; Cuadra y Veloso, 2010; Zubieta y Susinos, 1992). La satisfacción laboral del profesorado tiene implicaciones en su desempeño docente, por lo que evaluarla adquiere gran importancia para procesos de mejora en centros educativos.

En este contexto puede depender de dos grupos de variables o dimensiones, ambientales y personales, subrayando que los elementos relacionados con dichas dimensiones pueden influir en la satisfacción o la insatisfacción de los docentes. Las variables ambientales agrupan el clima laboral en el centro y el aula (Chaplain, 2008; Klassen y Chiu, 2010), la variabilidad de la actividad laboral, el salario (Butt et al., 2005; Crossman y Harris, 2006), las posibilidades de participación que ofrece la dirección y la aceptación por parte de la comunidad educativa (Greenglass y Burke,

2003). Las variables personales están relacionadas con la edad y el género (Fraser y Hodge, 2000; Klassen y Chiu 2010; Liu y Ramsey, 2008).

Compartimos con Tejero y Fernández (2009) que la satisfacción laboral es el estado emocional del docente o lo que es lo mismo, la vinculación afectiva al desempeño profesional realizado, ya sea de manera global o en distintas esferas del trabajo (Belkeman, 2004).

La satisfacción profesional de los docentes se compone de una serie de componentes estructurales que permiten la regulación de acciones pedagógicas y el comportamiento en las tareas del profesorado, y es un elemento crucial en la satisfacción en el trabajo docente (Caprara, Barbaranelli, Steca, y Malone, 2006). Por lo tanto, a través de los procesos de formación, todos los aspectos que pueden favorecer la satisfacción profesional se deben tomar en cuenta, entre los que se destacan: el compromiso de la enseñanza (Feather y Rauter, 2004), las condiciones de trabajo, la realización personal a través del desempeño y el bienestar profesional (Nir y Bogler, 2008; Vansteenkiste et al, 2007), el reconocimiento social de la profesión, las políticas educativas que normalizan el desempeño de los docentes y las posibilidades de desarrollo profesional (Hernández, 2014), y los beneficios de la autonomía en la enseñanza (Skaalvik y Skaalvik, 2014). Papel relevante que también tiene la institución, ya que “cuando la institución no respalda, no garantiza las condiciones de realización de la tarea, los sujetos la asumen en soledad, convirtiéndose en individualidades que tratan de afrontar la situación con sus propios medios y recursos personales”. (Alliaud y Vezub, 2012, p.950)

Destacar que estudios muestran que la satisfacción laboral del profesorado disminuye progresivamente conforme se avanza en antigüedad profesional, de ahí que los profesores más veteranos manifiestan estar menos a gusto con su trabajo actual y con la vida en el centro docente que sus colegas de menor antigüedad (Nieto y López-Martín, 2015). E incluso estudios realizados indican la diferencia entre el profesorado según el nivel donde desarrollan la actividad docente. Nieto y Riveiro (2006) los profesores de secundaria aparecen menos satisfechos que los de infantil y primaria. Guerrero (2003) en la universidad los niveles de insatisfacción vienen dados, entre otros, por la falta de apoyo de los compañeros, la indefinición de tareas y por la carencia de recursos.

2. Modelos y dimensiones vinculados a la satisfacción docente

Byrne (1999) valida un modelo configurado por cinco dimensiones que condicionan y conforman la satisfacción o insatisfacción docente, pudiendo generar situaciones de estrés laboral y, en determinados casos, burnout en el profesional docente. Según el autor, los factores del modelo hacen referencia al *conflicto de rol*, la *ambigüedad del rol*, la *carga de trabajo*, la *labor en el aula*, la *autonomía*, y la *relación con compañeros*. Siguiendo a este autor, el *conflicto de rol* es la dimensión vinculada a las expectativas de lo que debería ser el trabajo que realiza y lo que es en realidad; la *ambigüedad del rol*, está relacionada con la ausencia de protocolos o información clara respecto a su práctica docente, desde el punto de vista administrativo y educativo, en aspectos legales y responsabilidades; la *carga de trabajo*: es la dimensión relacionada con la cantidad y dificultad de trabajo que pueden llevar a cabo los docentes; la *labor en el aula* refiere

aspectos que influyen en la satisfacción del docente por la situación que vive en la clase, como el clima, el rendimiento y comportamiento del alumnado y la disciplina; *autonomía* es la dimensión que hace referencia a la libertad para tomar decisiones en situaciones de trabajo relacionadas con la práctica docente y el trabajo administrativo de los docentes; por último, la dimensión *relación con compañeros*, está asociada a la calidad de las interacciones que los docentes tienen con sus compañeros y la dirección del centro.

Respecto a estas últimas dimensiones, *autonomía y relaciones*, Evans (1999) señala que las interacciones entre el profesional docente y la dirección del centro son un factor determinante en la satisfacción de los docentes, incrementándose el bienestar a medida que los profesionales tenían más participación y estaban más implicados en la gestión, con el consiguiente aumento de la autonomía. Gran parte de la investigación que ha analizado la satisfacción del profesional docente, las motivaciones y el estrés laboral (Brunetti, 2001), tienen un nexo común en la necesidad de *autonomía* en el trabajo. Los estudios reflejan que la motivación del profesor y la autonomía están relacionadas con la satisfacción en el trabajo y la tensión a la que están sometidos. Los docentes con motivaciones intrínsecas están más satisfechos con sus empleos y son los que menos estrés experimentan (Davis y Wilson, 2000).

Pearson y Moomaw (2006), conscientes de la importancia de la autonomía en la satisfacción docente desarrollan una escala (*Teaching Autonomy Scale*) para poder evaluar los cuatro factores planteados en su modelo: selección de actividades y materiales, normas de conducta en el aula, la instrucción de planificación y secuenciación de la enseñanza y, por último, toma de decisiones personales en el trabajo.

La Escala de Satisfacción Laboral-Versión para Profesores (ESL-VP; Anaya y Suárez, 2007) evalúa la satisfacción laboral en torno a treinta y dos ítems que configuran cinco dimensiones denominadas: *diseño del trabajo*, es la dimensión relacionada con la satisfacción laboral en cuanto a la participación de los docentes en la planificación y diseño de objetivos relacionados con su trabajo; *condiciones de vida asociadas al trabajo*, es la dimensión vinculada con la situación del contexto escolar, la seguridad y los recursos existentes; *realización personal* conforma la dimensión que hace referencia a la percepción laboral positiva de los docentes en torno a su puesto de trabajo y el sentimiento de adecuación al mismo; *promoción y superiores*, evalúa las posibilidades de ascenso y la equidad de las personas con rango superior; por último, *salario* es la dimensión que establece la relación entre satisfacción y el sueldo recibido.

La escala de Anaya y Suárez combina dimensiones relacionadas con motivaciones extrínsecas, como el *salario* o la *promoción y superiores*, con dimensiones relativas a la motivación intrínseca, reflejadas en el *diseño del trabajo* o la *realización personal*. La satisfacción en el trabajo y la motivación pueden estar influidas por factores externos tales como el salario, la política educativa y las reformas o las condiciones en las que los docentes desempeñan su labor, aunque estos factores no son tan determinantes como los factores del contexto más inmediato al profesorado, relativos a las relaciones o su actividad en el aula (Evans, 1998).

La percepción de la adecuación al puesto de trabajo y el bienestar en el desempeño realizado forman parte de la concepción sobre las competencias que posee el profesorado, siendo un factor de satisfacción que influye en la eficacia en situaciones de estrés del profesorado (Dick y Wagner, 2001; Friedman, 2003). Los docentes que

poseen una percepción positiva sobre sus capacidades y desempeño tienen una satisfacción laboral mayor y contrarrestan mejor las situaciones problemáticas que aquellos que no se perciben como buenos profesores (Salanova, Cifre, Grau, Llorens y Martínez, 2005).

Satisfacción en el trabajo, a pesar de que es un factor estudiado, requiere una mayor profundización (Anaya y Suárez, 2007), siendo una de las prioridades para diseñar escalas con propiedades psicométricas. Por lo tanto, en esta escala, nos hemos centrado en proporcionar datos que sirvan para asumir la responsabilidad apropiada, como, medidas eficaces pertinentes para su mejora. Este instrumento no se basa en un modelo único, pero reúne los constructos relevantes e ítems de diferentes escalas. El punto común entre los estudios son las dimensiones de las relaciones personales, la planificación, la autonomía y el trabajo en clase (Anaya y Suárez, 2007; Brunetti, 2001; Byrne, 1999; Evans, 1999; Pearson y Moomaw, 2006). Las relaciones personales es sinónimo de la satisfacción con la interacción entre profesores dentro de la escuela (Skaalvik y Skaalvik, 2011). Planificación se refiere a la satisfacción con la programación y ejecución de actividades en la escuela y en casa (Nir y Bogler, 2008). La autonomía se centra en la satisfacción relacionada a la independencia en el desempeño de su trabajo (Skaalvik y Skaalvik, 2014). Por último, se refiere a la satisfacción con la actividad en el aula (Klassen y Chiu, 2010). Por lo tanto, el objetivo de este trabajo es elaborar una escala de satisfacción docente explorando si los datos recogidos con el instrumento, basado en factores comunes propuestos en la literatura científica, son fiables y válidos.

3. Metodología

3.1 Muestra

La muestra del estudio que describimos está constituida por 215 profesores de ambos sexos y de diferentes centros de una Comunidad Autónoma de España (islas de Gran Canaria y Lanzarote), que imparten docencia en la etapa de educación primaria, lo que representa el 43,26% del total de profesores que componen el claustro de los centros seleccionados.

3.2 Instrumento

El cuestionario para evaluar la satisfacción del profesional docente en la Enseñanza Primaria (ESEP) propuesto por el grupo de FODIP (2013), aplicado en el curso 2013-2014, consta de 18 ítems tipo Likert de 1 (nada satisfecho) a 5 (muy satisfecho) desarrollados con el objetivo de indagar el nivel de satisfacción de los profesionales en educación primaria. El análisis de la profesión docente obliga a explorar su actividad profesional desde los diferentes componentes estructurales que la conforman, es decir, las características de los sujetos que la realizan (su preparación, intereses, motivos), las condiciones en que tiene lugar el proyecto educativo, los métodos, las vías y los medios que se emplean para su desarrollo, así como las particularidades de los estudiantes en los que recae la acción pedagógica de los profesores y las políticas curriculares que enmarcan este proceso.

3.3 Procedimiento

En primer lugar calculamos los estadísticos descriptivos, media, desviación típica, asimetría y curtosis, de cada ítem. A continuación, para determinar la estructura factorial de la escala utilizamos la técnica multivariada, modelo de ecuaciones

estructurales exploratorio (MESE), desarrollado por Asparouhov y Muthen (2009). La ventaja principal de esta técnica es combinar el análisis factorial exploratorio (AFE) con el análisis factorial confirmatorio (AFC), además no requiere que el peso factorial de los ítems sea cero en los factores no correspondientes, por lo que el cálculo de los índices de ajuste y de las correlaciones entre variables latentes es más preciso (Asparouhov y Muthén, 2009; Marsh et al., 2009).

Respecto al método de rotación utilizado, siguiendo las recomendaciones de Asparouhov y Muthen (2009) para situaciones en las que se conoce poco de la estructura factorial a estudiar utilizamos una rotación geomin. Más específicamente, utilizamos el geomin oblícuo, ya que los métodos que implican rotaciones oblicuas, en las ciencias sociales, muestran relaciones entre factores más cercanas a la realidad (Brown, 2006; Schmitt, 2011).

En cuanto al método de estimación utilizado, al estar la escala formada por ítems tipo Likert, las variables observables son categóricas ordinales (Flora y Curran, 2004) y, por lo tanto, para estimar el valor de los parámetros y los índices de ajuste es más preciso utilizar un método de estimación que no requiere normalidad multivariada (Schmitt, 2011). Por ello, utilizamos el método de mínimos cuadrados ponderados ajustado por la media y la varianza (WLSMW).

Además el hecho de que los profesores estén agrupados por colegios viola el supuesto de independencia, esto puede inflar el valor de χ^2 e infraestimar los errores típicos (Stapleton, 2006). Para corregir esto, los parámetros fueron estimados maximizando una función logarítmica ponderada y los errores típicos utilizando un estimador tipo *sandwich* (Muthén y Muthén, 1998-2014). Para analizar el ajuste del modelo al patrón de los datos utilizamos la prueba de χ^2 , la ratio $\chi^2/g.l.$, el índice RMSEA y su intervalo con un nivel de cofianza del 90%, el índice de Tucker-Lewis (TLI), el índice de ajuste comparativo (CFI) y el residuo ponderado cuadrático medio (WRMR).

Para decidir el número de factores seguimos varios criterios, en primer lugar el sentido teórico de los mismos, en segundo lugar, siguiendo las recomendaciones de Costello y Osborne (2005), se debe elegir una estructura factorial en la que haya más de tres ítems por factor y con el menor número de pesos factoriales cruzados, por lo que en este estudio se prestó atención al número de factores con al menos tres ítems significativos en el factor esperado y que fueran significativos, con un nivel de confianza del 95%, únicamente en uno o dos factores y, por último, se atendió a los índices de ajuste, teniendo en cuenta que seguir este criterio puede llevar a aceptar más factores de los necesarios (Hayashi, Bentler y Yuan, 2007). Una vez determinado el número de factores, para evitar pesos factoriales cruzados se descartaron los ítems con pesos superiores a .35 en dos o más factores, luego se volvió a analizar la estructura factorial, así hasta que no hubiera ítems con pesos cruzados superiores a .35.

A continuación, se analizaron las evidencias de fiabilidad, para ello utilizamos el omega (ω) de McDonald (1999) a partir de la matriz de correlaciones policóricas, ya que este índice no requiere que el peso factorial sea igual para todos los ítems (Yang y Green, 2010) ni que los datos sean continuos (Elosua y Zumbo, 2008), al igual que con el alfa de Cronbach con valores por encima de .85 se considera que la escala es fiable. Los valores perdidos se estimaron utilizando un método imputación multivariante mediante ecuaciones encadenadas a través de la librería *mice 2.9* (Buuren y Groothuis-Oudshoorn, 2011) del *software R* (R Core Development Team, 2014). Para calcular los estadísticos descriptivos se utilizó el *software R*; para el MESE, el *Mplus 7.11* (Muthén

y Muthén, 1998-2014) y, por último, para estimar el omega de McDonald se utilizó la librería psych (Revelle, 2013) del programa estadístico R.

4. Resultados

4.1 Análisis descriptivos

En primer lugar se calcularon los estadísticos descriptivos de cada ítem. Como se puede observar en la Tabla 1, la media osciló entre 3.023 para el ítem 18 y 4.171 para el ítem 10, la desviación típica entre .711 y 1.123, ítems 15 y 18 respectivamente. En cuanto a la asimetría y la curtosis, los valores estuvieron por debajo de 2 en valores absolutos, es decir, valores bajos.

Tabla 1. Media, desviación típica y pesos factoriales de los ítems finales

	MEDIA	D.T.	P. ₁	A. ₂	PR. ₃
Planificación					
1. Satisfacción del profesorado para implicarse en el proyecto educativo de la escuela	3,89	,74	,39	,28	,26
2. Satisfacción del profesorado para participar en actividades extraescolares	3,80	,91	,79	-,01	,02
3. Satisfacción del profesorado para participar en actividades externas establecidas por la escuela	3,71	,92	,90	,01	-,01
Autonomía					
4. Satisfacción del profesorado para tomar iniciativas propias	3,94	,87	,12	,45	,25
5. Satisfacción del profesorado para trabajar en equipo con los colegas	4,00	,84	-,14	,82	-,01
6. Satisfacción del profesorado para relacionarse con otros colectivos de la escuela (comunidad, familia, etc.)	3,88	,84	-,01	,64	,23
7. Satisfacción del profesorado para la realización del plan de acción tutorial establecido por la escuela	3,80	,84	,04	,70	,14
8. Satisfacción del profesorado para implicarse en los procesos de comunicación de la escuela	3,73	,86	-,03	,90	,06
9. Satisfacción del profesorado para formar parte del equipo directivo u órganos de gestión de la escuela	3,57	1,09	,20	,52	-,08
Desempeño					
10. Satisfacción para la realización de su labor docente	4,09	,71	,02	,10	,71
11. Satisfacción en cuanto al proceso de enseñanza-aprendizaje en el aula	4,03	,71	-,03	,00	,98

Notas. 1: Planificación; 2: Autonomía; 3: Profesional.

Fuente: Elaboración propia.

4.2 Estructura factorial

En la Tabla 2 se puede observar que en la estructura factorial de 4 factores coincide el número de factores con tres o más ítems significativos en un factor y que no son significativos en más de uno o dos factores. Respecto a los índices de ajuste, se puede observar que a mayor número de factores, mejor ajuste, datos que están en consonancia con los resultados de Hayashi et al. (2007). Asparouhov y Muthen (2009) han observado, con puntuaciones continuas, que si añadir un factor disminuye el SRMR en .001 o menos, no merece la pena mantener la estructura con más factores. En este estudio, al ser los ítems variables categóricas se atendió al valor del RMSEA y así, a

raíz de los datos, podríamos descartar la opción de más de cinco factores. Finalmente, atendiendo a la discriminación de los factores y al valor de RMSEA se optó por la estructura de 4 factores. A continuación se eliminaron cuatro ítems por tener pesos superiores a .35 en dos o más factores, seguidamente se analizó la estructura factorial sin estos ítems y observamos que tres ítems tenían pesos superiores a .35 en dos factores, así que los descartamos y en el siguiente análisis de la estructura factorial se observó que un factor quedaba solo con un ítem, por lo que se decidió por una estructura factorial de tres factores, donde el primer factor, planificación, determinado por 3 ítems; el segundo factor, autonomía, formado por 6 ítems; y, finalmente, el tercer factor, desempeño docente, formado, únicamente, por 2 ítems. Esta estructura factorial cuenta con adecuados índices de ajuste $\chi^2(214, 33) = 109,15$ ($p = 0,00$), RMSEA = 0,10 [0,08; 0,10] CFI = 0,96, TLI = 0,91, and WRMR = 0,59 y con pesos factoriales en sus respectivos factores que oscilaron entre 0,392 y 0,976 (Ver Tabla 1).

Tabla 2. Número de factores con tres ítems significativos (95%) en un factor y que no son significativos en más de uno o dos factores e índices de ajuste en función del número de factores en cada extracción

FACTORES	FACT/ÍTEMS	X ²	G.L.	RMSEA	90%	CFI	TLI	WRMR
1	1	396.026	135	.093	.083 - .104	.905	.893	1.291
2	2	30.309	118	.083	.072 - .095	.934	.914	.966
3	2	236.337	102	.077	.064 - .090	.951	.927	.748
4	4	181.945	87	.700	.056 - .084	.966	.939	.575
5	3	122.621	73	.055	.038 - .072	.982	.962	.406
6	4	103.560	60	.057	.038 - .075	.984	.960	.351

Notas: Fact/Ítems =Número de factores que tienen al menos 3 ítems significativos en el factor que más pesan y no son significativos en más de dos factores (NC = 95%).

Fuente: Elaboracion propia.

4.3 Fiabilidad

Los resultados mostraron una fiabilidad adecuada tanto de la escala global de 11 ítems, $\omega_T=0,95$, como de las 3 dimensiones: planificación ($\omega=0,87$), autonomía ($\omega=0,89$) y desempeño docente ($\omega=0,87$).

5. Discusión y conclusiones

Como se puede observar en la introducción, las escalas sobre satisfacción laboral tienen distintas dimensiones aunque algunas de ellas se reiteran a lo largo de la literatura *ad hoc*, si no en su continente, sí en su contenido.

El objetivo de este estudio ha sido elaborar una escala de satisfacción docente, examinado las propiedades psicométricas de la escala desarrollada por el grupo FODIP (2013). En primer lugar se analizó la estructura factorial de la escala y, a continuación, se exploró la fiabilidad de la misma. Los resultados mostraron que la estructura factorial más apropiada es la de tres factores (planificación, autonomía y desempeño docente), al proporcionar un mejor ajuste, observando adecuadas evidencias de validez y fiabilidad.

Así, la dimensión planificación, con una adecuada fiabilidad ($\omega=0,87$), guarda relación con el denominado diseño del trabajo, en la escala de Anaya y Suárez (ESL-VP), una dimensión estrechamente relacionada con la participación, que otorga protagonismo a la implicación del profesorado en el diseño de los objetivos relativos a su encargo docente. La dimensión autonomía, en sus distintas concepciones, presentada sola o

acompañada de algún otro componente, constituye una de las categorías más analizada (Zubieta y Susinos, 1992; Evans, 1999; Byrne, 1999; Brunetti, 2001; Pearson y Moomaw, 2006). En este estudio, esta dimensión también presenta una fiabilidad ($\omega=.89$).

Por último, la dimensión desempeño docente refiere al trabajo del profesorado en el aula, en el desarrollo del proceso de enseñanza-aprendizaje. Es en relación con esta dimensión, en la esfera más próxima a su quehacer cotidiano como docente, donde el profesorado muestra un mayor grado de satisfacción. Byrne (1999) contempla la labor en el aula como una de las cinco dimensiones del modelo que propone; Anaya y Suárez (2007) incluyen la realización personal como una dimensión que alude a una percepción positiva del desempeño laboral del profesorado, existiendo evidencias sobre la relación entre una percepción positiva del desempeño laboral y un mejor manejo del estrés (Dick y Wagner, 2001; Friedman, 2003). Igualmente, los efectos del afrontamiento de situaciones problemáticas se ven amortiguados cuando existe una satisfacción con el desempeño profesional del profesorado (Salanova, Cifre, Grau, Llorens y Martínez, 2005).

Como puede apreciarse, las tres dimensiones de la escala encuentran correlatos similares en modelos anteriores y, aunque como ya señalamos, el número de dimensiones estimado por los distintos investigadores es variado, a juzgar por las referencias al respecto, una escala corta y de fácil aplicación puede ayudar a realizar una primera y sencilla aproximación a la satisfacción docente en los centros de educación primaria.

La escala resultante constituye un instrumento rápido y fácil en su manejo que puede contribuir a la realización de un primer diagnóstico del funcionamiento de un centro al evaluar la satisfacción de los docentes, ya que posee unas adecuadas propiedades psicométricas.

Como conclusión final, destacar que la escala constituye un instrumento rápido y fácil en su manejo y, al poseer unas adecuadas propiedades psicométricas, puede contribuir a la realización de un primer diagnóstico del funcionamiento de un centro al evaluar la satisfacción de los docentes.

5.1 Recomendaciones

En esta investigación tan solo se ha realizado un análisis preliminar de la escala diseñada por el grupo FODIP, por lo tanto, en futuras investigaciones sería recomendable, analizar las propiedades psicométricas de la escala con una muestra más extensa, si bien es cierto que se ha observado que se puede trabajar adecuadamente con muestras alrededor de los 200 sujetos como se ha reflejado en otros trabajos (Albaladejo-Blázquez, Ferrer-Cascales, Reig-Ferrer y Fernández-Pascual, 2013; Ruiz, Pardo y San Martín, 2010).

Por último, Fraser y Hodge (2000), plantean que variables personales como la edad puedan influir en la satisfacción laboral de los docentes, por lo tanto sería interesante analizar si los profesores con experiencia docente interpretan el cuestionario de forma similar a profesores con menos años de docencia.

Referencias

- Albaladejo-Blázquez, N., Ferrer-Cascales, R., Reig-Ferrer, A. y Fernández-Pascual, M. (2013). Does school violence occur in preschool and primary education? An assessment and management proposal. *Anales de Psicología*, 29(3), 1060–1069. doi:10.6018/analeps.29.3.158431.
- Alliaud, A. y Vezub, L. (2012). El oficio de enseñar. Sobre el quehacer, el saber y el sentir de docentes argentinos. *Revista Diálogo Educativo*, 12(37), 927-952. doi:10.7213/dialogo.educ.7211
- Anaya, D. y Suárez, J. M. (2007). Job satisfaction of Spanish teachers in preschool, primary and secondary education levels. A study within the national scope. *Revista de Educación*, 344, 217-243.
- Asparouhov, T. y Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling*, 16(3), 397-438. doi:10.1080/10705510903008204.
- Belkeman, S. (2004). Job Satisfaction. *CQ Weekly*, 62(40), 2420-2423.
- Brown, B. M. (2006) *Confirmatory factor analysis for applied research*. Nueva York: Guilford.
- Brunetti, G. (2001). Why do they teach? A study of job satisfaction among long-term high school teachers. *Teacher Education Quarterly*, 28(2), 49-74.
- Butt, G., Lance, A., Fielding, A., Gunter, H., Rayner, S. y Thomas, H. (2005). Teacher job satisfaction: Lessons from the TSW pathfinder project. *School Leadership and Management*, 25(5), 455–471. doi:10.1080/13634230500340807.
- Buuren, S. y Groothuis-Oudshoorn, K. (2011). MICE: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45(3), 1–67. doi:10.18637/jss.v045.i03.
- Byrne, B. (1999). The nomological network of teacher burnout: A literature review and empirically validated model. En R. Vandenberghe y A. Huberman (Eds.), *Understanding and preventing teacher burnout* (pp. 15-37). Cambridge: Cambridge University Press.
- Caprara, G., Barbaranelli, C., Steca, P. y Malone, P. (2006). Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level. *Journal of School Psychology*, 44(6), 473–490. doi:10.1016/j.jsp.2006.09.001.
- Carr, A. (2007). *Psicología Positiva. La ciencia de la felicidad*. Barcelona: Paidós.
- Chaplain, R. P. (2008). Stress and psychological distress among trainee secondary teachers in England. *Educational Psychology*, 28(2), 195–209. doi:10.1080/01443410701491858.
- Costello, A. y Osborne, W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment Research y Evaluation*, 10(4), 1-9.
- Crossman, A. y Harris, P. (2006). Job satisfaction of secondary school teachers. *Educational Management Administration and Leadership*, 34(1) 29–46. doi:10.1177/1741143206059538.
- Cuadra, A. y Veloso, C. (2010). Degree of supervision like moderating variable between the leadership and satisfaction, motivation and organizational climate. *Revista Chilena de Ingeniería*, 18(1), 15-25. doi:10.4067/s0718-33052010000100003.
- Cuscó, M. B. I. y Wells, J. G. (2009). La satisfacción del profesor de Educación Física. *Educación Física y Deporte*, 27(2), 27-35.
- Davis, J. y Wilson, S. (2000). Principals' efforts to empower teachers: Effects on teacher motivation and job satisfaction and stress. *The Clearing House*, 73(6), 349–357. doi:10.1080/00098650009599442.

- De Pablos, J., González, T. y González, A. (2008). El bienestar emocional del profesorado en los centros TIC como factor de innovación educativa. *Revista Latinoamericana de Tecnología Educativa*, 7(2), 45-55.
- Dick, R. y Wagner, U. (2001). Stress and strain in teaching: A structural equation approach. *British Journal of Educational Psychology*, 71(2), 243-259. doi:10.1348/000709901158505.
- Elousa P. y Zumbo, B. (2008). Coefficients of reliability for scales of categorical tidy response. *Psicothema*, 20, 896-901.
- Evans, L. (1998). *Teacher morale, job satisfaction and motivation*. Londres: Chapman.
- Evans, L. (1999). *Managing to motivate: A guide for school leaders*. Londres: Cassell.
- Feather, N. y Rauter, K. (2004). Organizational citizenship behaviors in relation to job status, job insecurity, organizational commitment and identification, job satisfaction, and work values. *Journal of Occupational and Organizational Psychology*, 77(1), 81-94. doi:10.1348/096317904322915928.
- Flora D. y Curran P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, 9(4), 466-491. doi:10.1037/1082-989x.9.4.466.
- Fraser, J. y Hodge, M. (2000). Job satisfaction in higher education: Examining gender in professional work settings. *Sociological Inquiry*, 70(2), 172-178. doi:10.1111/j.1475-682x.2000.tb00904.x.
- Friedman, I. (2003). Self-efficacy and burnout in teaching: The importance of interpersonal relations efficacy. *Social Psychology of Education*, 6(3), 191-215.
- Guerrero, E. (2003). Análisis pormenorizado de los grados de *burnout* y técnicas de afrontamiento del estrés docente en profesorado universitario. *Anales de Psicología*, 19(1), 145-158.
- Greenglass, E. y Burke, R. (2003). Teacher stress. En M. Dollard, H. R. Winefield y A. H. Winefield (Eds.), *Occupational stress in the service professions* (pp. 213-236.). Nueva York: Taylor and Francis.
- Hayashi, K., Bentler, P. y Yuan, K. (2007). On the likelihood ratio test for the number of factors in exploratory factor analysis. *Structural Equation Modeling*, 14(3), 505-526. doi:10.1080/10705510701301891.
- Hernández, A. (2014). La "Satisfacción Profesional": premisa o resultado del desarrollo profesional del profesorado. En T. Mentado y L. Castro (Coord.), *Satisfacción Profesional del profesorado de Educación Primaria. VI Seminario Internacional del Observatorio Internacional de la Profesión Docente* (pp. 13-25). Barcelona: Universidad de Barcelona.
- Hoppock, R. (1935.) *Job satisfaction*. Nueva York: Arno Press.
- Klassen, R. y Chiu, M. (2010). Effects on teachers' self-efficacy and job satisfaction: Teacher gender, years of experience, and job stress. *Journal of Educational Psychology*, 102(3), 741-756. doi:10.1037/a0019237.
- Liu, X. y Ramsey, J. (2008). Teachers' job satisfaction: Analyses of the Teacher Follow-Up Survey in the United States for 2000-2001. *Teaching and Teacher Education*, 24(5), 1173-1184. doi:10.1016/j.tate.2006.11.010.
- Marsh, H., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A. y Trautwein, U. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of University Teaching. *Structural Equation Modeling*, 16(3), 439-476. doi:10.1080/10705510903008220.

- McDonald, R. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Muthén, L. y Muthén B. (1998-2014). *Mplus User's Guide*. Los Ángeles, CA: Muthén & Muthén.
- Nieto, D. A. y López-Martín, E. (2015). Satisfacción laboral del profesorado de educación secundaria. *Revista de Investigación Educativa*, 33(2), 435-452. doi:10.6018/rie.33.2.202841.
- Nieto, D. A. y Riveiro, J. M. S. (2006). La satisfacción laboral de los profesores en función de la etapa educativa, del género y de la antigüedad profesional. *Revista de investigación educativa*, 24(2), 521-556.
- Nir, A. y Bogler, R. (2008). The antecedents of teacher satisfaction with professional development programs. *Teaching and Teacher Education*, 24(2), 377-386. doi:10.1016/j.tate.2007.03.002.
- Pearson, L. y Moomaw, W. (2006). Continuing validation of the teaching autonomy scale. *The Journal of Educational Research*, 100(1), 44-51. doi:10.3200/joer.100.1.44-51.
- Peiteado, M. G. (2013). Los estilos de enseñanza y aprendizaje como soporte de la actividad docente. *Journal of Learning Styles*, 6(11), 70.
- R Development Core Team. (2008). *R: A language and environment for statistical computing*. Viena: R Foundation for Statistical Computing.
- Revelle, W. (2013). *Psych: Procedures for personality and psychological research*. Evanston, IL: Northwestern University.
- Ruiz, M. A., Pardo, A. y San Martín, R. (2010). Structural equation models. *Papeles del Psicólogo*, 31(1), 34-45.
- Salanova, M., Cifre, E., Grau, R., Llorens, S. y Martínez, M. (2005). Precedents of the autoefficiency in teachers and university students: A causal model. *Revista de Psicología del Trabajo y de las Organizaciones*, 21(1-2), 159-176.
- Schmitt T. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment*, 29(4), 304-321. doi:10.1177/0734282911406653.
- Skaalvik, E. y Skaalvik, S. (2011). Teacher job satisfaction and motivation to leave the teaching profession: Relations with school context, feeling of belonging, and emotional exhaustion. *Teaching and Teacher Education*, 27(6), 1029-1038. doi:10.1016/j.tate.2011.04.001.
- Skaalvik, E. y Skaalvik, S. (2014). Teacher self-efficacy and perceived autonomy: Relations with teacher engagement, job satisfaction, and emotional exhaustion. *Psychological Reports*, 114(1), 68-77. doi:10.2466/14.02.pr0.114k14w0.
- Stapleton, L. (2006). An assessment of practical solutions for structural equation modeling with complex sample data. *Structural Equation Modeling*, 13(1), 28-58. doi:10.1207/s15328007sem1301_2.
- Tejero, C. M. y Fernández, M. J. (2009). Measurement of the labor satisfaction in the school direction. *Revista Electrónica de Investigación y Evaluación Educativa*, 15(2), 1-15.
- Vansteenkiste, M., Neyrinck, B., Niemiec, C. P., Soenens, B., De Witte, H. y Van den Broeck, A. (2007). On the relations among work value orientations, psychological need satisfaction and job outcomes: A self-determination theory approach. *Journal of Occupational y Organizational Psychology*, 80(2), 251-277. doi:10.1348/096317906x111024.

Yang, Y. y Green, S. (2010). A note on structural equation modeling estimates of reliability. *Structural Equation Modeling: A Multidisciplinary Journal*, 17(1), 66-81. doi:10.1080/10705510903438963.

Zubieta, J. y Susinos, T. (1992). *Las satisfacciones e insatisfacciones de los enseñantes*. Madrid: CIDE.

Evaluación Psicopedagógica e Intervención Sociocomunicativa en Niños con Trastorno de Asperger. Análisis de un Caso

Psychological Assessment and Socio-communicative Intervention for Children with Asperger Syndrome. A Case Analysis

Cecilia Latorre Cosculluela* y Miguel Puyuelo Sanclemente

Universidad de Zaragoza

El proceso de evaluación de necesidades sociocomunicativas de los niños con trastorno de Asperger debe estar adecuadamente justificado con el fin de definir las líneas de intervención en contextos naturales más oportunas, aprovechando las competencias para fomentar las habilidades más deficitarias de esta población. En este trabajo se pretende llevar a cabo una evaluación psicopedagógica de un niño con síndrome de Asperger para diseñar un trabajo de intervención en el área social y comunicativa apropiado, y que tiene como finalidad última la adaptación social del niño al entorno en que se encuentra inmerso. Se adoptó para ello el método cualitativo de estudio de caso con un enfoque interpretativo que posibilitara una comprensión profunda del caso en cuestión. De los datos recogidos se derivan una serie de necesidades educativas relacionadas con la conducta adaptativa del niño en la escuela y el establecimiento de relaciones sociales. Se proponen, además, algunas orientaciones de intervención social y comunicativa para potenciar un comportamiento funcional en el contexto escolar. De igual modo que los rasgos que caracterizan a la población de niños en general difieren significativamente entre sí, el caso aquí analizado es un claro ejemplo de la heterogeneidad que se detecta entre los niños con TEA en lo que a necesidades sociales y comunicativas se refiere. La evaluación de estas dimensiones se considera de fundamental importancia, siendo competencia de los profesionales del campo de la psicología y la educación la acomodación de las pautas establecidas para estos casos, a las particularidades de cada niño.

Palabras clave: Síndrome de Asperger, Necesidades sociocomunicativas, Evaluación psicopedagógica, Intervención social y comunicativa.

The assessment process of socio-communicative needs of children with Asperger syndrome must be properly justified in order to set the most appropriate and naturalistic lines of intervention, taking advantage of competences to promoting affected skills of this population. This paper aims to carry out a psychological assessment of a child with Asperger Syndrome in order to design an appropriate social-communicative intervention, whose ultimate goal is the social adaptation of the child to the setting in which he is included. For that it was taken the qualitative case study method with an interpretive approach that enables a deeper understanding of the case. All collected data illustrate a set of educational needs related to adaptive behavior of the child in school and establishing social relationships. Some social and communicative guidelines to promote the functional behavior of the child in the school context are also proposed. Just as defining characteristics of the population of children in general differ significantly from each other, the analyzed case here is a clear example of the diversity that is found among children with ASD in terms of social and communication needs are concerned. The assessment of these dimensions is considered a critical aspect. Educational and psychology professionals have the responsibility to adapt the guidelines that are established for such cases to the particularities of these children.

Keywords: Asperger Syndrome, Socio-communicative needs, Psychopedagogical assessment, Social and communicative intervention.

1. Introducción¹

Considerando la versión actualmente más consultada del Manual Diagnóstico y Estadístico de los Trastornos Mentales publicada en 2002, el DSM-IV-TR, el Trastorno de Asperger se caracteriza por una inhabilidad grave y permanente en la dimensión de interacción social y la manifestación de pautas de conducta e intereses repetitivos y limitados. El trastorno puede repercutir de manera clínicamente significativa en la competencia social y laboral, así como en otras áreas esenciales de la actividad funcional de las personas (American Psychiatric Association, 2002).

Recientemente, las continuas discrepancias entre los investigadores del tema han originado numerosas opiniones acerca de la validez de la etiqueta del síndrome. Con el fin de esclarecer esta situación, en la última versión de este manual publicada en 2013 viene establecido un continuo sintomatológico en el que se incluyen desde individuos con leves síntomas autistas hasta otros con problemas más severos, todos ellos englobados bajo la categoría de Trastorno de Espectro Autista (en adelante, TEA) (APA, 2013, citado por APA, 2014). Cabe señalar, no obstante, que este concepto ha sido recientemente cuestionado por otros autores (Tsai, 2013; Perry, 2014), quienes manifiestan su oposición a este continuo defendiendo la utilidad clínica que tienen algunos diagnósticos anteriormente utilizados en este espectro para orientar a los clínicos en la determinación del pronóstico y los planes de tratamiento.

En nuestros días, bien es sabido que uno de los desafíos más complejos que el sistema educativo debe afrontar, se relaciona con ofrecer una respuesta adecuada a las necesidades educativas de su alumnado (Lledó, Lledó y Pomares, 2011). En concreto, la atención educativa a los niños con síndrome de Asperger se considera un proceso de fundamental importancia, y tiene la finalidad de contribuir a una mejora del proceso educativo de los mismos.

Dentro del conjunto de necesidades que presenta esta población en el ámbito educativo, las alteraciones cualitativas de la interacción social se consideran un aspecto definitorio de la misma (Sansosti, 2010). Más concretamente, se manifiestan por una serie de dificultades para formar amistades verdaderas, fracaso en el uso y comprensión de pautas no verbales de comunicación, y ausencia de reciprocidad social y emocional. Por otro lado, y aunque uno de los criterios diagnósticos que se establecen en el DSM-IV-TR para el síndrome de Asperger es la ausencia de un retraso del lenguaje significativo, los datos de diferentes estudios (Attwood, 2009; Etchepareborda, 2001) sugieren que el retardo precoz del lenguaje no debería ser un criterio de exclusión sino de inclusión. Finalmente, los altos niveles de rigidez mental y conductual que presentan estos niños se manifiestan a través de patrones comportamentales que conducen a una falta de interés por compartir sus conocimientos de manera recíproca (Dahle y Gargiulo, 2004). Tomando en consideración la nueva clasificación que aparece en el DSM-5, los niños que anteriormente se diagnosticaban con trastorno de Asperger, pasarían ahora a

¹ Este artículo ha estado financiado por fondos del Departamento de Ciencia y Tecnología del Gobierno de Aragón al grupo de investigación consolidado en Psicología: Cognición, Lenguaje, Aprendizaje y Ciclo Vital. Aspectos Evolutivos y Educativos.

considerarse dentro de ese abanico de TEA, más concretamente de alto funcionamiento.

Los criterios que se acaban de describir suponen, en la mayoría de los casos, un obstáculo que puede interferir en la interacción social con sus iguales. El hándicap más significativo de este espectro se localiza en el funcionamiento social (Adams, Green, Gilchrist y Cox, 2002). Incluso en aquellos casos de niños con un grado de funcionalidad no significativamente afectado, las dificultades en el área social siguen constituyéndose como una de las necesidades educativas más características (Paul, Miles, Chuba y Volkmar, 2009).

La población de niños con TEA de alto funcionamiento necesita que, desde los contextos educativos, se les enseñe cómo funcionan las rutinas y los diferentes modos de establecer interacciones con otros compañeros, de tal manera que logren manejar su propia conducta y comprender las reglas sociales que rigen las conversaciones (Barquero, 2007). A causa de su poca habilidad para deducir, de forma intuitiva, los pensamientos, planes e intenciones de otras personas, los niños con síndrome de Asperger y, en general, con TEA, requieren además que se les aporten los recursos necesarios para comprender conductas, anticipar reacciones y relacionarse apropiadamente con los demás. En muchas ocasiones, y de acuerdo con Barquero (2007), el problema no es la poca interacción social, sino más bien la eficacia de ésta.

Puesto que son niños que asimilan y retienen mejor la información que les llega por la vía visual, las estrategias que adopten esta fuente de entrada de datos serán las más adecuadas para comprender mejor a este alumnado y lograr así la consecución de objetivos educativos (Lledó et al., 2011). En el área social, y utilizando el método de enseñanza que se acaba de mencionar, será necesario que estos niños se encuentren con tareas que persigan, a grandes rasgos, la modificación de patrones de conductas no adecuados sustituyéndolos por comportamientos sociales coherentes con cada situación.

El perfil comunicativo de estos niños de alto funcionamiento varía desde algunos que presentan marcadas alteraciones cualitativas y cuantitativas en su habilidad para comprender y transmitir la información, hasta otros que cuentan con un desarrollo del lenguaje verbal apropiado para su edad (Shriver, Allen y Mathews, 1999). En estos últimos casos siguen detectándose ciertas peculiaridades en su proceso comunicativo que interfieren en su desempeño social y lingüístico y que resultan condicionantes directos de su funcionamiento sociocomunicativo tanto en el ámbito escolar como en otros contextos de interacción.

En lo que respecta a los aspectos formales del lenguaje, autores como González Carbajal (2002) señalan una conservación de los mismos en los niños de alto funcionamiento, pudiendo presentar no obstante un “patrón específico de desarrollo de habilidades morfosintácticas y semánticas fuertemente asociado a este trastorno” (Martín Borreguero, 2005, p.117). En términos de comportamiento pragmático, estos niños ven alterada su comprensión del lenguaje figurativo, el sarcasmo y el humor, la prosodia, la expresión facial y otras formas de lenguaje no verbal que intervienen en la función comunicativa (Williams, Goldstein, Kojkowski y Minschew, 2008). Consecuentemente, aparece en estos niños la necesidad de aprender habilidades de relación social que, desde una perspectiva pragmática del lenguaje, se concibe como la capacidad de actuación social (González Carbajal, 2002).

Tal y como se observa en el estudio de Paul et al. (2009), los aspectos conversacionales más significativamente afectados y que deberían ser tomados en consideración en todo planteamiento de intervención, son la cantidad de información proporcionada en la conversación para satisfacer las necesidades del oyente, la capacidad para mantener y para responder a las señales sociales. Otras áreas incluyen la capacidad para iniciar y mantener un discurso espontáneo e inferir los estados mentales de los demás. Nuevamente, los apoyos visuales serán los que permitan al niño aprender, a través de una enseñanza explícita, a expresar sus ideas y explicarse correctamente, a respetar los turnos de palabra y a comprender las expresiones faciales ligadas a las emisiones verbales de otras personas.

En el área comunicativa, otra de las necesidades educativas de la población de niños con síndrome de Asperger (o con TEA de alto funcionamiento, si atendemos a la nueva clasificación del DSM-5) se relaciona con su capacidad narrativa. Las intervenciones que se planteen presentarán un trabajo sobre este aspecto, ayudando al niño a no emitir información irrelevante, a seguir una secuencia temporal y a elaborar narraciones en las que se incluyan referencias a los estados mentales de los personajes (González Carbajal, 2002). Diversas investigaciones, entre ellas la de Losh y Capps (2003) han encontrado relaciones entre el rendimiento en tareas de falsa creencia de niños con TEA y la frecuencia de descripciones de estados mentales y afectivos de los personajes de las historias narrativas. Precisamente por ello, el trabajo de intervención con estos niños incluirá, necesariamente, tareas en las que se vea implicada la capacidad de atribución de estados mentales.

Analizando las consecuencias educativas que el síndrome tiene para el niño que lo padece, destaca su falta de interés para participar en los juegos sociales, tendiendo por ende a un juego repetitivo y aislado (Frontera, 2010; Lledó et al. 2011). Cuando se incorporan a la escuela comienzan a aparecer conflictos de adaptación al aula con alteraciones conductuales, una tendencia a evitar interacciones sociales espontáneas y dificultades para mantener conversaciones sencillas. Si bien manifiestan la necesidad de ser aceptados socialmente por sus compañeros, en ausencia de tales interacciones estos niños presentan sentimientos de aislamiento y depresión (Bonete, Vives, Fernández, Calero y García, 2010). Por tanto, desconocen cómo interactuar adecuadamente debido a su ajustada comprensión social y emocional.

Desde el ámbito psicoeducativo, la necesidad de seguir planteando el análisis de las necesidades que aparecen en niños con síndrome de Asperger (o TEA de alto funcionamiento) es absolutamente evidente, con la finalidad de dar cabida a una atención adaptada a las características de cada niño y de su entorno. Así, un trabajo explícito sobre el ámbito social y comunicativo puede suscitar mejoras considerables siempre y cuando, en la línea de lo expuesto por Monfort (2009), los contenidos se extraigan de las necesidades reales del niño y se introduzcan después en su contexto.

La pretensión de este estudio es presentar a los educadores, psicopedagogos y demás profesionales que trabajan a diario con alumnos con TEA, una ejemplificación del proceso de evaluación psicopedagógica y de la propuesta de intervención más oportuna dadas las necesidades educativas de un caso particular. En un primer momento se consideró necesario analizar el perfil de habilidades sociales y lingüísticas del niño, de tal manera que los datos recogidos sirvieran como punto de partida para el diseño de una intervención desde el contexto de escolaridad ordinaria. La integración de toda la información permitió una descripción del funcionamiento actual del niño y de sus

necesidades así como la determinación de las estrategias necesarias para estimular su desarrollo y aprendizaje.

2. Método

El presente trabajo de investigación se elaboró a través de una metodología cualitativa utilizando el estudio de caso, presentando éste un carácter descriptivo-interpretativo, pues pretende describir los diferentes factores que ejercen influencia en el fenómeno objeto de estudio (Martínez Carazo, 2006). La intención final del mismo es generalizar desde un punto de vista analítico, es decir, utilizar este trabajo para ilustrar y representar una teoría ya existente.

La elección del caso se justifica por dos criterios fundamentales, a saber: en primer lugar, por el aliciente de conocer, específicamente, las necesidades sociales y comunicativas de un niño que todavía no ha alcanzado la etapa de Educación Primaria; y, en segundo lugar, por la posibilidad de acceso tanto al niño en cuestión como al entorno escolar en que se encuentra inmerso.

2.1. Unidad de análisis

La unidad de análisis es situacional ($n=1$), es decir, un único sujeto con unas características peculiares alrededor del cual se organiza la recogida de datos y las orientaciones de intervención. El caso de estudio es un niño escolarizado, desde hace tres años, en un colegio público de enseñanza ordinaria, y que actualmente se encuentra en el nivel de 3º de Educación Infantil. Hasta la fecha, no ha sufrido enfermedades destacables ni se encuentran bajo tratamiento médico. Además, su evolución respecto a los aprendizajes escolares es propia de su edad.

Para la elaboración de este estudio, tanto familia como escuela debían presentar las autorizaciones correspondientes. Previamente, la tutora del aula del niño informó de que, aparentemente, el niño presentaba una capacidad cognitiva adecuada a su edad. La maestra advierte también una serie de comportamientos extraños con el resto de compañeros de la clase que se traducen en ciertas dificultades para acercarse a interactuar con ellos. Además de presentar ecolalias y un contacto ocular episódico, tiene un interés obsesivo por las letras y los números.

2.2. Instrumentos

Con el propósito de asegurar un mayor rigor y objetividad en los resultados del estudio, fueron varios los instrumentos de recogida de datos seleccionados (Stake, 1998). Para analizar el área social, se utilizó una escala de evaluación de la conducta adaptativa. Por otro lado, la detección de necesidades derivadas de la dimensión lingüística y comunicativa se realizó por medio de un instrumento estandarizado, una lista de chequeo y un registro de observación durante la resolución de tareas que ponían en juego las capacidades narrativas del niño.

Para identificar las dimensiones lingüísticas que podrían estar afectadas se utilizó el *PLON-R* (Prueba de Lenguaje Oral de Navarra) (Aguinaga, Armentia, Fraile, Olangua y Uriz, 2004), un instrumento elaborado para detectar problemas lingüísticos a edades tempranas. La tutora del aula del niño cumplimentó una de las grandes escalas de evaluación de la conducta adaptativa en el actual panorama internacional, el *ABAS-II* (Oakland y Harrison, 2013). Permite una evaluación directa de los principales déficits que presentan los niños con TEA, incluidas las habilidades sociales y de comunicación.

Se analizaron además los resultados obtenidos tras la administración del CCC-2 (Children's Communication Checklist) (Bishop, 2003) a la tutora del niño.

En cuanto a la evaluación de habilidades narrativas, se utilizaron tres series de secuencias con apoyo visual en base a las que el niño debía crear un relato a través del lenguaje oral. Este análisis se justifica, de acuerdo con Baixauli, Roselló y Miranda (2004), por las dificultades observadas en niños con TEA para utilizar términos mentalistas, pudiendo además crear historias narrativas con una cantidad pobre de información y una desestructuración de las ideas del discurso.

3. Resultados

La información recabada se analizó por medio de la técnica de la triangulación de fuentes, un principio que aporta racionalidad, de tal manera que las categorías examinadas son: el síndrome de Asperger, las necesidades sociales y comunicativas, y el caso propiamente dicho. La fiabilidad de la investigación se aseguró contando con una base de datos que sustenta el estudio de casos, de tal modo que, con los datos recogidos, un nuevo investigador pudiera repetir el estudio y llegar a conclusiones similares.

Una de las necesidades educativas más evidentes que se detectan tras la cumplimentación del ABAS-II se localiza en la dimensión de interacción social del niño, siendo de hecho el ámbito más significativamente afectado que aparece en esta escala. Causada probablemente por estas dificultades primarias en el establecimiento de interacciones sociales, se aprecia además la necesidad de plantear una intervención en lo que respecta a la competencia comunicativa del niño y su funcionamiento en el ámbito escolar. Así, y en términos generales, el resultado que se desprende en cuanto a la conducta adaptativa general (CAG) sitúa al niño en un nivel "bajo".

La información obtenida en el cuestionario CCC-2 ofrece datos relevantes sobre su desarrollo lingüístico. Unos resultados que orientan acerca del planteamiento alrededor del cual debe girar la intervención psicoeducativa en el ámbito escolar, y que parte de las necesidades detectadas en las escalas de uso del contexto, relaciones sociales, intereses e iniciación inadecuada. Si bien en sintaxis y semántica no se aprecian déficits relevantes, el nivel de funcionamiento está por debajo de lo que se esperaría para un niño de su edad.

Las puntuaciones extraídas en las distintas dimensiones lingüísticas que incluye el PLON-R sirven también de indicadores acerca de los posibles ámbitos hacia los que urge plantear el trabajo de intervención. En las dimensiones de "forma" y "contenido" del lenguaje, el niño se sitúa en un rango de "necesita mejorar", mientras que en la dimensión de "uso" correspondiente al ámbito pragmático, se encuentra ubicado en un nivel de "retraso". El ámbito pragmático se establece, por tanto, como el área prioritaria de intervención dentro de la competencia comunicativa del niño.

En cuanto a la producción narrativa utilizando las secuencias de imágenes, las observaciones y anotaciones realizadas señalan ciertos aspectos de la coherencia narrativa que se ven alterados durante la producción verbal y que se traducen en un conjunto de necesidades que orientarán el trabajo de intervención. Dado que no establece relaciones de significado entre los elementos que componen las secuencias narrativas, uno de los requisitos bajo los que se plantearán las posteriores tareas, se relaciona con la utilización de apoyos visuales que dirijan y orienten la elaboración de las historias. De acuerdo con los resultados obtenidos en esta prueba de carácter

cualitativo, este niño precisará además un trabajo explícito sobre atribución de estados mentales y emociones a los demás, pues sus dificultades de comprensión se incrementan cuando la historia adquiere un carácter más emocional.

Una vez examinada la información se presentan, en el siguiente espacio, las líneas de intervención en habilidades psicosociales y lingüísticas que, dadas las necesidades del niño, deberían guiar el trabajo en el centro educativo en que se encuentra escolarizado.

3.1. Intervención en habilidades psicosociales y lingüísticas

Para atender a las dificultades de niños con unas características similares a las del caso al que se refiere este trabajo (síndrome de Asperger o, de acuerdo con la nueva versión del DSM, TEA de alto funcionamiento), es necesario disponer de apoyo de pedagogía terapéutica orientado a ofrecerles la ayuda necesaria para comprender las diferentes rutinas, aceptar los cambios e interactuar con los compañeros. Durante la etapa de escolarización se requiere además apoyo logopédico con el fin de potenciar el desarrollo de los distintos ámbitos del lenguaje y, especialmente, de la comunicación (Frontera, 2010).

Tras conocer las necesidades educativas de este niño, lo cual supone un verdadero reto para muchos profesionales, se planteó la respuesta educativa encaminada a desarrollar al máximo sus habilidades sociales y comunicativas con vistas a la aplicación de los saberes adquiridos. Dada la gran diversidad existente dentro del espectro autista respecto al grado de severidad y el modo en que los síntomas se manifiestan, resultaría inapropiado aplicar orientaciones de intervención de carácter general sin llevar a cabo una individualización significativa (Frontera, 2012). Se diseñaron así una serie de orientaciones que se basarían en los siguientes principios de intervención que Lledó et al. (2011) proponen:

- ✓ Enseñar de forma explícita a través de instrucciones y apoyos visuales.
- ✓ Estructurar el entorno de aprendizaje para que se convierta en predecible.
- ✓ Anticipar los posibles cambios y proponer diferentes alternativas.

En lo que respecta al ámbito social, los objetivos que se establecen para este caso son, por un lado, apropiarse de las normas sociales que rigen las conversaciones y, por otro, comprender situaciones sociales específicas que encuentra confusas a través de una enseñanza explícita de habilidades de inicio, mantenimiento y finalización de conversaciones ajustadas a cada situación. Se establecen, por ende, las siguientes orientaciones específicas para la intervención en el área psicosocial:

- ✓ Una herramienta de gran utilidad para enseñar al niño a enfrentarse a situaciones sociales son los guiones (scripts) sociales. Consisten, de acuerdo con Frontera (2010, p.220), en “una descripción explícita de la secuencia de pasos a seguir en una situación específica, indicando el tipo de comportamiento que se espera de ellos”.
- ✓ Es conveniente mostrarle, de la manera más objetiva y perceptible posible, las maneras de llevar a cabo una interacción social exitosa trabajando cuestiones como “por qué sucede algo”, “cuáles son las consecuencias que se producen en el interlocutor cuando no se respeta su turno de palabra” y demás aspectos relacionados con situaciones de interacción. En este sentido, las “conversaciones en forma de historieta” de Carol Gray (1998) pueden resultar una herramienta de gran utilidad a partir de las que podemos personalizar las situaciones que queremos enseñarle. Teniendo en cuenta los resultados

obtenidos en el ABAS-II, este material puede resultar de gran utilidad para trabajar en el niño aquellos aspectos que se encuentran más deteriorados como, por ejemplo, las actividades del día a día en el aula y la escuela.

- ✓ Role-playing: este tipo de metodología de trabajo con el niño supondrá, de acuerdo con autores como De la Iglesia y Olivar (2005) la toma de conciencia de “qué” y “cómo” decir y hacer en un buen número de situaciones sociales. Estos aprendizajes ofrecerán al niño una base sobre la que ir construyendo un repertorio de comportamientos acordes a cada circunstancia específica.
- ✓ Para trabajar la inferencia de estados mentales y lograr una mejor comprensión de las situaciones sociales, pueden utilizarse dibujos en los que aparezcan imágenes de distintas situaciones en las que los personajes se ven inmersos. Un material de gran utilidad es el elaborado por Monfort y Monfort (2002), donde aparecen diversas tareas destinadas a estimular los procesos mentalistas de niños con dificultades. Algunos ejemplos son los que aparecen las Figuras 1 y 2.



Figura 1. Ilustraciones extraídas del material "En la Mente"

Fuente: Monfort y Monfort (2002).



Figura 2. Viñeta para trabajar emociones secundarias

Fuente: Rodríguez, García, Gorriz y Regal (2002).

- ✓ Una vez se hayan trabajado los aspectos anteriormente comentados, conviene generalizar estos aprendizajes en contextos de interacción con otros niños, en forma de pequeño o gran grupo. El entrenamiento sistemático en habilidades sociales ayuda a desarrollar habilidades comunicativas y resolver problemas en situaciones de interacción social (Griffin, Griffin, Fitch, Albera y Gingras, 2006). Pero no olvidemos que “el aprendizaje será más efectivo si se le proporcionan al niño experiencias significativas que se produzcan dentro de contextos de interacciones sociales” (González Carbajal, 2002, p.97).
- ✓ Rutinas diarias: El orden secuencial de los eventos y el establecimiento de rutinas estables fomentarán en el niño la sensación de orden y la habilidad para predecir los acontecimientos cotidianos (Frontera, 2012), ayudando así al niño a evitar situaciones de estrés y ansiedad.

Por otro lado, el objetivo inicial de cualquier intervención en el desarrollo comunicativo va dirigido a “promover el aprendizaje de formas adecuadas de regulación de otras personas, alternativas al empleo de formas de comunicación inadecuadas” (Gortázar, 2003, p.183). En este caso, la meta de la intervención sobre los aspectos lingüísticos y comunicativos es lograr que el niño pueda comunicarse mejor con sus interlocutores habituales. Por ello, la intervención logopédica en niños con TEA de alto funcionamiento se centra fundamentalmente en la enseñanza del uso social adecuado del lenguaje. Las orientaciones específicas más adecuadas para este caso son las que a continuación aparecen expuestas:

- ✓ Los juegos de rol y las dramatizaciones constituyen una ventajosa herramienta para trabajar diversos aspectos de la pragmática, el discurso y la conversación, ámbitos en los que este niño manifiesta considerables dificultades. Esta metodología permitirá además trabajar los movimientos corporales, los gestos, la prosodia y los sentimientos propios y ajenos.
- ✓ Para trabajar específicamente las habilidades lingüísticas pragmáticas, se puede utilizar el apoyo visual de secuencias que conforman una historia y que el niño deberá construir y narrar utilizando el lenguaje oral. En este caso, la intervención en este aspecto es esencial, pues las dificultades en la narración pueden limitar el acceso a la interacción con otros afectando significativamente a la competencia sociocomunicativa del niño. Nuevamente, el material “En la Mente” de Monfort y Monfort (2002) puede emplearse también para trabajar el área pragmática del lenguaje del niño, más concretamente la comprensión de preguntas y de situaciones de mentira y broma.
- ✓ Las tareas de lectura pueden resultar un instrumento muy interesante para favorecer el discurso narrativo de este niño. De acuerdo con González Carbajal (2002), y a partir de la lectura de una pequeña historia utilizando los apoyos necesarios, se puede dialogar sobre el contenido, seleccionando los aspectos relevantes del mismo y extrayendo las ideas principales.
- ✓ Las conversaciones en forma de historieta de las que hablábamos en el apartado anterior han demostrado también tener una gran eficacia para trabajar habilidades comunicativas. Usando dibujos, “bocadillos” de pensamiento, símbolos y diferentes colores, se pueden mostrar y enseñar al niño diferentes aspectos tales como el turno de palabra, los cambios en el tema de conversación, las clarificaciones y la evitación de comentarios irrelevantes.
- ✓ Las actividades de comunicación referencial resultan una herramienta muy útil para trabajar situaciones comunicativas (González Carbajal, 2002). El niño tiene que contar o describir algo a otra persona sin ver el objeto en cuestión a describir.
- ✓ Una vez que dejemos atrás la enseñanza explícita de estas situaciones y normas sociales, pasaremos a trabajarlas en un contexto más funcional de interacción: practicándolo con el adulto en una situación simulada. Durante esta actividad el especialista reforzará las actitudes adecuadas del niño y ofrecerá modelos para sustituir aquellos intentos erróneos de establecimiento de interacción. Como soporte a este proceso pueden utilizarse apoyos visuales de fotografías, dibujos y tarjetas, que fomentarán la atención y comprensión del niño durante los intercambios comunicativos (Frontera, 2012).
- ✓ El nivel más complejo para esta actividad es el que se refiere a la puesta en práctica de las habilidades aprendidas en un contexto real de interacción con sus iguales. Se aprovecharán situaciones en las que un grupo de niños estén jugando y se animará al niño a que se acerque a ellos y ponga en práctica lo que se ha trabajado.

Todas las actuaciones que se han descrito en este trabajo se apoyan en las necesidades psicoeducativas que tiene este niño con síndrome de Asperger, y tienen como objetivo prioritario una intervención que conduzca al desarrollo y optimización de habilidades de interacción social para la mejora de las relaciones interpersonales. Además, la función de los educadores, psicopedagogos, orientadores y logopedas es crucial en este proceso, debiendo contar todos ellos con una formación competente para colaborar en

el proceso de evaluación, en la búsqueda de alternativas de respuesta y en la valoración del proceso educativo (Frontera, 2010).

3. Discusión

Pese a ser un niño que sigue un patrón normalizado de evolución, presenta las necesidades educativas específicas del diagnóstico de espectro autista. Consecuentemente, y de igual modo que se encuentra en las investigaciones de estudio de caso de Jurado y Bernal (2011) y Vives y Ascanio (2012), las áreas más afectadas son las relacionadas con la interacción social, el lenguaje y la comunicación. En esta línea, los resultados que Saulnier y Klin (2007) obtienen en su estudio con niños con síndrome de Asperger y Autismo de Alto Funcionamiento señalan, igual que lo encontrado en este trabajo, la necesidad de realizar evaluaciones de las habilidades comunicativas y sociales de cada niño para determinar del grado de funcionalidad adaptativa en contextos de la vida real.

En el caso escogido para la elaboración de este estudio, y coincidiendo con los resultados del estudio de Jurado y Bernal (2011), las necesidades más notables derivadas de la evaluación psicopedagógica se localizan en el ámbito pragmático del lenguaje, que se encuentra estrechamente relacionado con el área social. Por el contrario, las dificultades de comprensión social y lingüística descubiertas en este niño discrepan con los hallazgos del trabajo de Lin y Chiang (2014), quienes sugieren que no existe una diferencia significativa entre niños con trastorno de Asperger y niños con desarrollo típico en cuanto a comprensión se refiere.

Por otro lado, las puntuaciones obtenidas en el ABAS-II confirman estas necesidades sociocomunicativas. De acuerdo con los resultados del estudio de validez del ABAS-II (Oakland y Harrison, 2013, p.159) con una muestra de niños autistas, los resultados aquí encontrados están en consonancia con los de estos autores, aunque en el caso del que se ocupa este estudio se obtiene un grado de funcionalidad de conducta adaptativa general (CAG) más elevado que en el grupo de estos autores.

Las revisiones llevadas a cabo por autores como Johnson et al. (2007) también señalan un conjunto de características que están presentes en niños diagnosticados dentro del espectro autista y que deben ser interpretadas, desde el contexto escolar, como necesidades a tener en consideración por los profesionales que van a trabajar con estos niños. En concordancia con lo encontrado en este caso, estos autores determinan la aparición de dificultades conversacionales y de interacción comunicativa que deben ser valoradas en la intervención posterior. En cuanto a las necesidades derivadas del uso del contexto, en otros trabajos como el de Loukusa et al. (2007) también se concluye la pertinencia de ofrecer a estos niños apoyo en la interpretación de emisiones comunicativas en función de la situación en que se encuentran inmersos.

De manera complementaria, el análisis de las habilidades narrativas ha permitido descubrir que la inferencia de aspectos que no se manifiestan de manera explícita en la historia, es deficitaria en este niño lo que, según Whyte y Nelson (2015), es un aspecto relativamente frecuente en niños con TEA. También en la revisión de estudios sobre inferencia pragmática en individuos con Asperger y AAF que Loukusa y Moilanen (2009) llevan a cabo, se demuestra para ambas condiciones la presencia de dificultades y, en consecuencia, la necesidad de iniciar una intervención, en aquellas tareas donde se ponen en práctica habilidades de inferencia pragmática y entendimiento de

intencionalidad, ofreciendo como posible explicación al respecto los déficits detectados en Teoría de la Mente que tradicionalmente se han atribuido a esta población. Por estas razones, algunos de los contenidos planteados en el proceso de intervención tienen por finalidad potenciar, en la medida de lo posible, todos estos aspectos.

Los resultados que se acaban de exponer deben interpretarse con cautela pues, siguiendo lo expuesto por Van der Paelt, Warreyn y Roeyers (2014), la edad cronológica es un significativo predictor del desarrollo del lenguaje que, además, se encuentra estrechamente relacionado con la competencia sociocomunicativa de cada niño. Por ello, cabe la posibilidad de que, con el transcurso de los años de escolaridad, estos niños con TEA de alto funcionamiento evolucionen considerablemente en todas estas dimensiones en las que ahora se localizan necesidades específicas de apoyo psicoeducativo.

Si bien una investigación como la propuesta puede limitar la generalización de resultados, se constata la importancia de seguir planteando el análisis de las realidades que suponen la presencia de diversas necesidades y características en el ámbito escolar. De acuerdo con los resultados del estudio de Vives y Ascanio (2012), las intervenciones que parten del análisis de las necesidades sociocomunicativas de los niños tienen la virtud de lograr generalizar, en muchas ocasiones, otras conductas que no son tratadas directamente. Por ende, tan importante como planificar una o varias sesiones de evaluación de las competencias del niño, es el planteamiento de una intervención ajustada a sus necesidades. La adaptación de estas intervenciones puede tener, indudablemente, efectos claramente favorecedores para su desarrollo comunicativo y social, tal y como se muestra en el trabajo de Adams (2003) con niños con Trastorno Pragmático del Lenguaje.

Este enfoque, aunque comience con un trabajo individual, deberá ir orientado a la aplicación en contextos naturales y con interlocutores cotidianos, potenciando los entornos comunicativos y el papel activo del niño (González Carbajal, 2002). Desde el punto de vista de cualquier profesional responsable de implementar una intervención, lo fundamental es adaptar las orientaciones generales a las particularidades de cada niño y su entorno. Se enseña de manera explícita lo que de forma natural no se ha podido aprender, y se adapta esa enseñanza a cada individuo en particular (Monsalve, 2002), muestra de lo cual es el caso aquí estudiado. El alcance de este tipo de intervenciones puede comprobarse también en estudios como el de Millar, Light y Schlosser (2006) con personas con trastornos del desarrollo, encontrándose una mejora significativamente positiva en la producción comunicativa y lingüística.

El apoyo de los educadores, terapeutas y familia, así como el diseño de intervenciones adecuadas, puede ayudar a los niños con Trastorno de Asperger (y, en general, a los niños con TEA) a afrontar más eficazmente sus dificultades y funcionar de un modo más independiente en el contexto familiar, escolar y comunitario (Griffin et al. 2006). Por ende, es tarea de los profesionales especializados en el campo de la psicología y la educación acomodar las pautas de trabajo a las individualidades de cada niño. Este asesoramiento pedagógico incluido dentro de un proyecto inclusivo de intervención será el que, de acuerdo con Darretxe y Sepúlveda (2011), facilite a estos niños el acceso a los aprendizajes en las mejores condiciones.

Referencias

- Adams, C., Green, J., Gilchrist, A. y Cox, A. (2002). Conversational behaviour of children with Asperger syndrome and conduct disorder. *Journal of Child Psychology and Psychiatry*, 43(5), 679-690. doi:10.1111/1469-7610.00056.
- Adams, C. (2003). Intervention for developmental pragmatic language impairments. *Aula Abierta*, 82, 79-96.
- Aguinaga, G., Armentia, M., Fraile, A., Olangua, P. y Uriz, N. (2004). *Prueba de Lenguaje Oral de Navarra-Revisada (PLON-R)*. Pamplona: Fondo de Publicaciones del Gobierno de Navarra.
- American Psychiatric Association. (2002). Trastornos generalizados del desarrollo. En *DSM-IV-TR, Manual diagnóstico y estadístico de los trastornos mentales* (pp. 79-97). Barcelona: Masson.
- American Psychiatric Association. (2014). Trastornos generalizados del desarrollo. En *DSM-5, Manual diagnóstico y estadístico de los trastornos mentales* (pp. 125-148). Buenos Aires: Panamericana.
- Attwood, T. (2009). ¿Qué es el síndrome de Asperger? El diagnóstico. En *Guía del síndrome de Asperger* (pp. 19-89). Barcelona: Paidós.
- Barquero, M. (2007). *Síndrome de Asperger. Guía para padres de familia y educadores*. San José: Vicerrectoría de Acción Social, Universidad de Costa Rica.
- Bishop, D. V. M. (2003). *Children's Communication Checklist (CCC-2)*. Recuperado de <https://moodle.usal.es/>
- Baixauli, I., Roselló, B. y Miranda, A. (2004). Evaluación de las dificultades pragmáticas. Estudio de casos. *Revista de Neurología*, 38(1), 69-79.
- Bonete, S., Vives, M. C., Fernández, A., Calero, M. D. y García, M. B. (2010). Potencial de aprendizaje y habilidades sociales en escolares con el trastorno de Asperger. *Behavioral Psychology/Psicología Conductual*, 18(3), 473-490.
- Dahle, K. B. y Gargiulo, R. M. (2004). Understanding Asperger disorder: A primer for early childhood educators. *Early Childhood Education Journal*, 32(3), 199-203. doi:10.1023/b:ecj.0000048973.89175.a3.
- Darretxe, L. y Sepúlveda, L. (2011). Estrategias educativas para orientar las necesidades educativas de los estudiantes con síndrome de Asperger en aulas ordinarias. *Electronic Journal of Research in Educational Psychology*, 9(24), 869-892.
- De la Iglesia, M. y Olivar, J. S. (2005). Orientaciones para la intervención psicopedagógica en niños con síndrome de Asperger y autismo de alto nivel de funcionamiento cognoscitivo. *Puertas a la lectura*, 18, 70-83.
- Etchepareborda, M. C. (2001). Perfiles neurocognitivos del espectro autista. *Revista de Neurología Clínica*, 2(1), 175-192.
- Frontera, M. (2010). Cuarta propuesta: El desafío del desarrollo. Alumnos con síndrome de Asperger. Necesidades y respuesta educativa. En E. Arrimadas, S. Ramos y A. Díaz-Güemes (Coords.), *Desafíos de la diferencia en la escuela* (pp. 179-252). Madrid: Edelvives.
- Frontera, M. (2012). Intervención: Principios y programas psicoeducativos. En M. A. Martínez y J. L. Cuesta (Coords.), *Todo sobre el autismo* (pp. 221-270). Tarragona: Altaria.
- González Carbajal, A. (2002). Intervención comunicativa en autismo de alto funcionamiento y síndrome de Asperger. En J. Martos y M. Pérez Juliá (Coords.), *Autismo: Un enfoque orientado a la formación en logopedia* (pp. 85-106). Valencia: Nau Llibres.

- Gortázar, P. (2003). Intervención comunicativa en casos con autismo: Estado de la cuestión. Presentación de un caso clínico. En F. Alcántud (Coord.), *Intervención psicoeducativa en niños con Trastornos Generalizados del Desarrollo*. Madrid: Pirámide.
- Griffin, H. C., Griffin, L. W., Fitch, C. W., Albera, V. y Gingras, H. (2006). Educational interventions for individuals with Asperger syndrome. *Intervention in School and Clinic*, 41(3), 150-155. doi:10.1177/10534512060410030401.
- Johnson, C. P. y Myers, S. M. (2007). Identification and evaluation of children with Autism Spectrum Disorders. *Pediatrics*, 120(5), 1183-1215. doi:10.1542/peds.2007-2361.
- Jurado, P. y Bernal, D. (2011). El alumno afectado con síndrome de Asperger en el aula ordinaria. *Revista de Educación Inclusiva*, 4(2), 29-46.
- Lin, Y. H. y Chiang, H. M. (2014). Language comprehension of children with Asperger's disorder and children with autistic disorder. *Research in Autism Spectrum Disorders*, 8(7), 767-774. doi:10.1016/j.rasd.2014.03.018.
- Lledó, A., Lledó, G. y Pomares, J. (2011). Las necesidades del alumnado con síndrome de Asperger: orientaciones y propuestas desde una innovación inclusiva. En R. Roig Vila y C. Lavene (Eds.), *La práctica educativa en la sociedad de la información. Innovación a través de la investigación* (pp. 243-254). Alcoy: Marfil.
- Losh, M. y Capps, L. (2003). Narrative ability in high-functioning children with autism or Asperger's syndrome. *Journal of Autism Developmental Disorders*, 33(3), 239-251.
- Loukusa, S., Leinonen, E., Kuusikko, S., Jussila, K., Mattila, M. L., Ryder, N., Ebeling, H. y Moilanen, I. (2007). Use of context in pragmatic language comprehension by children with Asperger syndrome or high-functioning autism. *Journal of Autism Developmental Disorders*, 37(6), 1049-1059. doi:10.1007/s10803-006-0247-2.
- Loukusa, S. y Moilanen, I. (2009). Pragmatic inference abilities in individuals with Asperger syndrome or high-functioning autism. A review. *Research in Autism Spectrum Disorders*, 3(4), 890-904. doi:10.1016/j.rasd.2009.05.002.
- Martín Borreguero, P. (2005). Perfil lingüístico del individuo con síndrome de Asperger: Implicaciones para la investigación y la práctica clínica. *Revista de Neurología*, 41(1), 115-122.
- Martínez Carazo, P. C. (2006). El método de estudio de caso. Estrategia metodológica de la investigación científica. *Pensamiento y Gestión*, 20, 165-193.
- Millar, D. C., Light, J. C. y Schlosser, R. W. (2006). The impact of augmentative and alternative communication intervention on the speech production of individuals with developmental disabilities: A research review. *Journal of Speech, Language, and Hearing Research*, 49(2), 248-264. doi:10.1044/1092-4388(2006/021).
- Monfort, M. y Monfort, I. (2002). *En la mente: Un soporte gráfico para el entrenamiento de las habilidades pragmáticas en niños*. Madrid: Entha Ediciones.
- Monfort, M. (2009). Comunicación y lenguaje: Bidireccionalidad en la intervención en niños con trastorno de espectro autista. *Revista de Neurología*, 48(2), 53-56.
- Monsalve, C. (2002). Comunicación y lenguaje en autismo: Claves para el buen trabajo de un logopeda. En J. Martos y M. Pérez Juliá (Coords.), *Autismo: Un enfoque orientado a la formación en logopedia* (pp. 65-83). Valencia: Nau Llibres.
- Oakland, T. y Harrison, P. (2013). *ABAS-II. Sistema para la evaluación de la conducta adaptativa*. Madrid: TEA Ediciones.

- Paul, R., Miles, S., Chuba, H. y Volkmar, F. (2009). Conversational behaviors in youth with high-functioning ASD and Asperger syndrome. *Journal of Autism Developmental Disorders*, 39(1), 115-125. doi:10.1007/s10803-008-0607-1.
- Perry, R. (2014). Asperger's disorder on life support. *Journal of Autism Developmental Disorders*, 44(8), 2072-2073. doi:10.1007/s10803-014-2067-0.
- Rodríguez, N., García, E. M., Gorriz, A. B. y Regal, R. (2002). *¿Cómo se estudia el desarrollo de la mente?* Recuperado de <http://www.uji.es/bin/publ/edicions/jfif8/psi/20.pdf>
- Sansosti, F. J. (2010). Teaching social skills to children with autism spectrum disorders using tiers of support: A guide for school-based professionals. *Psychology in the Schools*, 47(3), 257-281. doi:10.1002/pits.20469.
- Saulnier C. A. y Klin, A. (2007). Brief report: Social and communication abilities and disabilities in higher functioning individuals with autism and Asperger syndrome. *Journal of Autism Developmental Disorders*, 37(4), 788-793. doi:10.1007/s10803-006-0288-6.
- Stake, R. E. (1998). La triangulación. En *Investigación con estudio de casos* (pp. 94-104). Madrid: Morata.
- Shriver, M. D., Allen, K. D. y Mathews, J. R. (1999). Effective assessment of the shared and unique characteristics of children with autism. *School Psychology Review*, 28(4), 538-558.
- Tsai, L. Y. (2013). Asperger's disorder will be back. *Journal of Autism Developmental Disorders*, 43(12), 2914-2942. doi:10.1007/s10803-013-1839-2.
- Van der Paelt, S., Warreyn, P. y Roeyers, H. (2014). Social-communicative abilities and language in preschoolers with autism spectrum disorders: Associations differ depending on language age. *Research in Autism Spectrum Disorders*, 8(5), 518-528. doi:10.1016/j.rasd.2014.01.010.
- Vives, M. C. y Ascanio, L. (2012). Resultados del tratamiento en un caso con trastorno de Asperger. *Behavioral Psychology*, 20(2), 401-419.
- Whyte, E. M. y Nelson, K. E. (2015). Trajectories of pragmatic and nonliteral language development in children with autism spectrum disorders. *Journal of Communication Disorders*, 54, 2-14. doi:10.1016/j.jcomdis.2015.01.001.
- Williams, D. L., Goldstein, G., Kojkowski, N. y Minshew, N. (2008). Do individuals with high functioning autism have the IQ profile associated with nonverbal learning disability? *Research in Autism Spectrum Disorders*, 2(2), 353-361. doi:10.1016/j.rasd.2007.08.005.

Una Noción de Evaluación: La Evaluación Formativa en el Marco de las Concepciones de Finlandia y la Ciudad de Buenos Aires

An Assessment Notion: Formative Assessment in the Frame of the Conceptions in Finland and the City of Buenos Aires

Celina Cardoner Zunino*

Universidad de San Andrés

El término evaluación ha sido históricamente una palabra que se utiliza frecuentemente en la práctica docente, con distintos sentidos. Existen diferencias respecto del modo de conceptualizarla que hacen que la evaluación de los aprendizajes sea una temática de difícil resolución. Este informe de investigación se propone identificar los argumentos generados en torno a la cultura evaluativa de la ciudad de Buenos Aires, comparándolos con las tendencias de evaluación que existen en Finlandia. A su vez, luego de afirmar que la evaluación formativa es indispensable para la mejora en la calidad de los aprendizajes, se identificará la presencia o ausencia de argumentos que se inscriban o no en el paradigma de la evaluación formativa. Por ello, trabajando en torno a los propósitos de evaluación que se presentan en los Diseños Curriculares vigentes de escuela media -el *National Core Curriculum for Upper Secondary Schools* y el Diseño Curricular de la Nueva Escuela Secundaria- se identificarán las rupturas, los vacíos y concepciones respecto de la evaluación que se presentan formalmente para evaluar los aprendizajes. Los resultados demuestran que ambos documentos incorporan la evaluación formativa como herramienta para conocer y trabajar en el aprendizaje del alumno pero demuestran diferencias significativas en la concepción de sujeto que hay por detrás de esa pedagogía. Las autoridades educativas de ambas regiones exponen el deseo de atender personalmente a los alumnos dándoles feedback formativo pero, sin embargo, le exigen distintas habilidades al futuro egresado.

Palabras clave: Aprendizaje significativo, Retroalimentación, Evaluación formativa, Aprendizaje.

The term evaluation has historically been a word that is often used in teaching, with different meanings. There are differences in how to conceptualize that make learning assessment a topic difficult to resolve. This research report aims to identify the arguments generated around the evaluation culture of the city of Buenos Aires, comparing assessment trends that exist in Finland. In turn, after saying that formative assessment is essential for improving the quality of learning, we will identify the presence or absence of arguments that sign in or not in the paradigm of formative assessment. Therefore, working around evaluation purposes served by existing middle school curricular designs -the *National Core Curriculum for Upper Secondary Schools* and *Curriculum Design of New School Secondary*- we will identify breaks, gaps and conceptions regarding the assessment to be formally presented to assess learning on the middle level. Through the reading of documents, the ideas of evaluation and speeches about the subject role will make the analysis more profound and complex. The results show that both documents incorporate formative assessment as a tool to learn and work in student learning but show significant differences in the subject role that is put into play behind this pedagogy. Education authorities in both regions exhibit a desire to personally attend training students giving feedback but, however, they require different skills to future graduates.

Keywords: Student achievement, Formative assessment, Feedback, Learning.

*Contacto: celinacardoner@gmail.com

issn: 1989-0397

www.rinace.net/riee/

<https://revistas.uam.es/riee>

Recibido: 7 de noviembre de 2015

1ª Evaluación: 20 de enero de 2016

Aceptado: 16 de febrero de 2016

1. Introducción

Desde que se creó el Programa Internacional para la Evaluación de Estudiantes (PISA), Argentina ha obtenido resultados que lo ubicaron entre los últimos puestos mientras que Finlandia siempre ha estado entre los primeros. Por lo tanto, el estado de la evaluación en la actualidad argentina introduce un marco complejo en el campo: la política educativa no se centra completamente en el alumno ni en las capacidades que los chicos deberían desarrollar en la escuela (habilidades intrapersonales, interpersonales y cognitivas). Desde el Ministerio de Educación, no sólo de la Ciudad Autónoma de Buenos Aires (CABA) sino también de Finlandia, se publicaron informes que describen la concepción de evaluación, la forma en la que se deberían evaluar los aprendizajes de los alumnos en la escuela media y las habilidades que se esperan de los alumnos egresados. Estos documentos son la puerta de entrada para hacer una primera aproximación a lo que se concibe como evaluación en las diferentes regiones. Ahora bien, partiendo del supuesto de que se plantean distintas estrategias, significados y silencios en torno a la evaluación de aprendizajes en las distintas regiones escogidas, este artículo contribuye en:

- ✓ La exploración de las diversas concepciones paradigmáticas de evaluación de los aprendizajes que predominan en las prácticas educacionales de Buenos Aires y Finlandia.
- ✓ La identificación de los modos y configuraciones que contribuyen y se inscriben en la noción de evaluación formativa en los distintos Diseños Curriculares de las regiones seleccionadas y el efecto que tienen en la constitución de un futuro egresado.

¿Cómo tiene que ser la evaluación de los aprendizajes para que el estado de esta cuestión no permanezca en decadencia? La respuesta se encuentra en los postulados de Scriven (1967), quien hace cincuenta años introdujo el término evaluación formativa. En los sistemas escolares actuales este tipo de evaluación debería ser de especial relevancia como dispositivo de pedagogía diferenciada. Se percibe una necesidad de individualizar el contenido, el ritmo y las modalidades de aprendizaje en función de objetivos definidos con claridad.

Por esto, la evaluación formativa es un quiebre, una ruptura al proponer dejar a un lado el grado de regulación grupal y trasladarlo a la individualización; busca una intervención acorde al proceso individual de cada alumno. De esta manera, el docente se enfrenta a un cambio abrupto en las dinámicas escolares en relación a la enseñanza, la administración de horarios, la organización de la clase, entre otras cosas.

Por su parte, la concepción cognitiva postula que el aprendizaje debe ser significativo y, para ello, el alumno debe interactuar con su entorno de manera tal que construya sus representaciones personales y debe realizar juicios de valor que le permitan tomar decisiones en base a ciertos parámetros de referencia. Es decir, la evaluación debe partir de los conocimientos y vivencias previas del sujeto para luego integrar el nuevo conocimiento y, de esa manera, generar aprendizaje significativo.

Es indispensable que en el proceso de evaluación se pongan en juego las capacidades del alumno, su comprensión, su análisis, sus actitudes, intereses y habilidades, para poder perfeccionar -a través de la retroalimentación- la acción educativa y el desarrollo integral del educando. Si la evaluación es formativa, entonces el aprendizaje será significativo.

Al analizar los distintos documentos provenientes de las dos regiones, el *National Core Curriculum for Upper Secondary Schools* y el Diseño Curricular de la Nueva Escuela Secundaria (DCNES), se identifican las distintas propuestas institucionales que desarrolló cada gobierno para su sistema de evaluación y se ponen en juego las diferentes estrategias, tendencias, aportes y objetivos en relación a la evaluación y a lo que se espera del alumno.

Así se establece -por ejemplo- en el DCNES, que se espera que “la evaluación tenga como foco principal el logro de la autonomía personal y académica de los estudiantes, pudiendo desarrollarse mediante procesos de evaluación, autoevaluación y co-evaluación” (Ministerio de Educación de la Ciudad de Buenos Aires, 2013, p. 145).

Por su parte, el *National Core Curriculum for Upper Secondary Schools* publicado por el *Finnish National Board of Education* plantea que la evaluación de los alumnos tiene como objetivo orientar y fomentar el aprendizaje y desarrollo de los estudiantes de diversas formas. El papel de la evaluación del aprendizaje de los estudiantes es proporcionarles información sobre sus avances y resultados de aprendizaje durante y tras la finalización de los estudios de enseñanza secundaria (Finnish National Board of Education, 2003).

A pesar de ser Finlandia parte de la Unión Europea y una región económicamente más activa que la Ciudad de Buenos Aires, forma parte de este análisis comparativo porque es un modelo de referencia dentro del mundo desarrollado. Es de utilidad conocer tanto las similitudes como las diferencias que rigen en la formalidad de los documentos para tener una mirada global de los distintos aterrizajes que tiene el concepto de evaluación en las regiones alrededor del mundo, especialmente de la región con mayor éxito educativo en los últimos años y en la región más representativa de la Argentina.

2. Fundamentación de los documentos a analizar

Los documentos a ser analizados son los diseños curriculares vigentes de escuela media de dos regiones diferentes. Por un lado, el *National Core Curriculum for Upper Secondary Schools* publicado por el *Finnish National Board of Education* en el año 2003 (está en proceso el nuevo diseño curricular que se va a lanzar en 2016); por otro lado, el Diseño Curricular de la Nueva Escuela Secundaria de la Ciudad de Buenos Aires publicado en el año 2013.

Se han elegido estos documentos por sobre otros porque, además de ser de circulación pública, son las propuestas formales que se presentan desde los ministerios y, por lo tanto, las propuestas que todas las instituciones deberían adquirir y poner en juego.

Ambos documentos presentan las dinámicas y lógicas institucionales que se esconden por detrás de los sistemas educativos y brindan una descripción exhaustiva de los contenidos a enseñar y las formas en las cuáles se distribuirán los contenidos a lo largo de la escolaridad secundaria. En tal sentido, estos documentos brindan herramientas para que las distintas instituciones planteen recorridos en los que se pongan en juego variadas formas de construcción, apropiación y reconstrucción de saberes para los estudiantes de nivel secundario. Además, proponen un marco común que permite reconocer y hacer lugar a la diversidad de trayectorias escolares, intereses y puntos de partida de los adolescentes.

Como fue planteado anteriormente, si bien Finlandia es un país que se encuentra en una región más desarrollada económicamente y posee diferencias culturales significativas,

es relevante hacer una comparación debido al impacto que tiene este sistema educativo: recibe regularmente las mejores calificaciones a escala mundial y está reconocido como un modelo de referencia dentro del mundo desarrollado. Los resultados obtenidos periódicamente en el Informe PISA, así como el elevado índice de educación, respaldan al modelo igualitario y gratuito de educación en Finlandia.

Es habitual encontrar noticias en los diarios, publicaciones en Facebook o charlas TED que mencionen lo exitoso y prestigioso que es estudiar en Finlandia. Este país tiene el mejor sistema educativo del mundo por el valor que se le da a la profesión docente, por el método educativo que tiene, por cómo son los centros educativos, por la cultura y la política educativa. Por ese motivo, y también teniendo en cuenta los últimos resultados de las pruebas PISA, resulta eficiente e importante reconocer cómo circula el discurso de la evaluación formativa alrededor del mundo y cuál es el impacto que tiene en la mejora de la calidad de los aprendizajes. Se ha optado por el Diseño Curricular de la CABA por la relevancia e impacto que tiene en la matrícula escolar, siendo una de las ciudades del país con mayor número de escuelas públicas.

Sabiendo que el tema a desarrollar es la evaluación de los aprendizajes, hay ciertas secciones que serán de especial relevancia por sobre las demás presentadas en los distintos documentos. En el caso del DCNES, el capítulo se titula: “La evaluación en la NES” mientras que en el diseño curricular finlandés se presenta bajo el nombre de: “*Assessment of student learning*”. Estas secciones exponen no sólo una concepción de evaluación particular sino también explican las herramientas puntuales a utilizar en el campo de la evaluación de los aprendizajes. Son estas las secciones que nos permitirán identificar la ausencia o presencia de argumentos sobre evaluación formativa.

De todas maneras, para conocer el perfil del egresado que se espera y relacionarlo con los propósitos de evaluación que promueve cada región, se tomarán los argumentos planteados en otras secciones de los documentos: las habilidades del educando, los objetivos de la escuela media, los contenidos y la cultura de enseñanza.

3. Metodología

En función de la naturaleza de los objetivos de la investigación y sabiendo que se intenta explorar el tipo de discurso que hay en torno a la evaluación y al perfil del egresado que se espera por parte de la CABA y Finlandia, la estrategia metodológica del estudio se inscribirá en una perspectiva cualitativa. Un estudio de esta índole nos permite recoger los discursos completos para proceder a la interpretación de los mismos en relación a la evaluación formativa, analizando las relaciones de significado que se producen con otras nociones.

Es decir, a través de un análisis documental de los diseños curriculares seleccionados -nuestras fuentes de datos-, se pondrán en juego, en primer lugar, los argumentos de Cowen (2000) para poder hacer una lectura global de la cuestión. Esta primera aproximación permite pensar en las educaciones comparadas y ayuda a explicar ciertos procesos sociales y educacionales complejos como lo es la evaluación de los aprendizajes.

En segundo lugar, suponiendo que la lectura global crea un discurso universal para las distintas regiones, se utilizarán los aportes de Popkewitz (2005) para profundizar y dar cuenta de qué manera “aterriza” el concepto de evaluación formativa tanto en la CABA

como en Finlandia; la noción de *traveling libraries* clarifica de qué manera se trasladan los supuestos alrededor del mundo.

Ahora bien, el proceso a llevarse a cabo tendrá en cuenta “hasta qué punto podemos aprender algo del valor práctico del estudio de sistemas educativos extranjeros” (Cowen, 2000, p. 13). Por ese motivo y dado que la educación está inserta en lo social, está contextualizada y relacionada con otros elementos, es que Cowen también nos ayudará a conocer de qué manera trabaja la noción de evaluación formativa en cada región de manera singular, en lo local.

Para abarcar todas las contribuciones propuestas al comienzo, se optará por complementar el discurso global de evaluación formativa que prima con el discurso local de cada una de las regiones escogidas. Entonces esta investigación será una herramienta que intente plantear que “el mundo es una mezcla de lo global y lo local” (Cowen, 2000, p. 20).

En una segunda instancia, utilizando la noción de *traveling libraries* (Popkewitz, 2005), se analizarán los documentos para ver de qué manera se reconstruyen y ensamblan los discursos de evaluación formativa en los distintos entornos culturales. Esta noción nos permite reconocer cómo las culturas locales se van adaptando y reorientando a las ideas del exterior en formas históricamente específicas. Las ideas educativas siempre han viajado por todo el mundo, sin embargo, hay que dar cuenta del proceso de recontextualización (Ball, 2002).

Por último, sabiendo que detrás de las estrategias de evaluación existe un objetivo por parte del docente que evalúa, se analizará la configuración del alumno al realizar las evaluaciones. ¿Cuál es el objetivo final de la evaluación? ¿Cómo tiene que responder el alumno? ¿Qué se espera del alumno?

A lo largo de todo el análisis de los documentos se intentará responder a estas cuestiones tomando los aportes de Hacking (2006), quien afirma que el estudio de ciertas ciencias crean tipos de personas que, en cierto sentido, no existían antes. Es decir, los distintos aterrizajes de la concepción de evaluación y las diferentes propuestas que se promueven por parte de los ministerios serán analizadas a luz del *making-up people*: esta noción proporciona una manera de hablar y de interpretar la evaluación de los aprendizajes y la manera en la que se pone en juego en el mundo. Este análisis se hará a su vez de la mano de Fenstermacher y Soltis (1998) que propondrán un marco teórico para encasillar el tipo de docentes que se presenta por parte de las distintas regiones.

4. Análisis de los documentos: Finlandia y CABA

4.1. Contextualización: “La evaluación de los aprendizajes”

La evaluación de los aprendizajes es una práctica social anclada en un contexto determinado: está conceptualizada e interpretada con significados distintos en las diferentes regiones; se hacen usos dispares, con fines e intenciones diversas, utilizando variedad de instrumentos, siguiendo principios y normas diferentes para dar a entender que, en su aplicación, sigue criterios de calidad.

Desde siempre, los sistemas de evaluación han perdurado en lo que conocemos como el sistema escolar. Lo interesante aquí es determinar de qué manera se inscribe un tipo de práctica -la evaluación formativa- en la contemporaneidad de nuestros tiempos, en el

contexto actual; es decir, desarmar el concepto de “evaluación” para desnaturalizar su condición hegemónica ya que se ha vinculado a la evaluación con la acreditación en la mayoría de los institutos escolares de la ciudad de Buenos Aires. Más aún, resaltar la inserción de esta práctica en lo que se define como modernidad.

Siguiendo la línea de pensamiento de Foucault (1994), la evaluación formativa puede caracterizarse como una práctica moderna; una práctica que permite el desarrollo de una actitud particular. Siendo la modernidad un modo particular de relacionarse con y frente a lo que sucede cotidianamente, esta práctica provee al alumno una manera de pensar y de sentir diferente ya que incorpora el ejercicio de trabajar sobre sí mismo, en su pensamiento crítico y en su relación con el mundo real.

Por lo tanto, al evaluar formativamente el aprendizaje del alumno, se pasa del conocimiento tradicional basado en el contenido, a la incorporación de nuevos conocimientos que forman parte de una novedad: el sujeto como actor activo y reflexivo, configurando una nueva práctica de evaluación, de un nuevo sistema educativo. Además, para poder incorporar esta actitud de modernidad, el individuo necesita ser consciente de que es objeto de una elaboración ardua y compleja que no se detiene.

Esa conciencia la aprende a través de la evaluación formativa porque esta práctica genera *life-long learners*, genera la actitud -en el alumno- de querer aprender constantemente, recibiendo feedback a lo largo de todo el desarrollo educativo para poder inventarse y elaborarse a sí mismo una y otra vez. Por este motivo, la evaluación debería ser considerada como un proceso y no como un suceso, y construirse en un medio y nunca en un fin; la evaluación es el proceso a través del cual criticamos lo que somos, lo que tenemos, lo que alcanzamos pero también a través del cual experimentamos la posibilidad de conducir y modificar esas críticas que actúan como límites (Foucault, 1994).

Esta actitud moderna implicaría lo que para Kant (1784) es salirse del estado de tutela. En el contexto de la evaluación formativa, se le permite al alumno operar por él mismo sobre sí mismo, sabiendo que “la vocación de todo hombre es a pensar por sí mismo” (Kant, 1784, p. 8). Por tanto, esta práctica busca dejar a un lado la racionalidad técnica y el aprendizaje memorístico para dar lugar a la evaluación formativa y el aprendizaje significativo, dándole al alumno cierto grado de libertad y autonomía de manera tal que pueda conducirse en las circunstancias diarias de la vida.

Como parte de la modernidad, la evaluación formativa atiende a los alumnos de manera personal, intentando reestructurar su conocimiento a partir de las actividades que llevan a cabo e incitándolos en el libre pensar para llegar a convertirlos en algo más que una máquina que repite: los convierte en alumnos capaces de construir su propio pensamiento. Esto es un ejemplo de que el protagonista de la modernidad es el sujeto-alumno; el foco está puesto fundamentalmente en los logros que él puede realizar por sí sólo y, más aún, la evaluación formativa evalúa la habilidad que tiene el alumno de poder inventarse a sí mismo y auto-regularse. La forma de regulación que ejerce la evaluación formativa supone intervenir necesariamente durante el período de tiempo dedicado a una unidad de formación, adaptando la enseñanza antes de que se califique al alumno o se decida su aprobación o desaprobación.

Por último, cabe señalar que la práctica de evaluación formativa y significativa surge en el contexto de la modernidad por el carácter reflexivo que requiere el proceso. Como parte de esta actitud moderna, el alumno, a través de la evaluación formativa y significativa, consigue diferenciarse de las modalidades tradicionales manteniéndose en contacto con lo que hace y reflexionando sobre su propia acción (Giddens, 1990). Es decir, consigue definir su ethos, entendiéndolo como el conjunto de rasgos y modos de comportamiento que conforman el carácter o la identidad del alumno.

De esta manera, el desempeño escolar de los individuos es examinado constantemente por los alumnos y, con ayuda de las retroalimentaciones de los docentes, reformado a la luz de la nueva información que se obtiene a partir de las prácticas de evaluación. Por tanto, “la modernidad está totalmente constituida por la aplicación del conocimiento reflexivo” (Giddens, 1990, p. 47).

Todo proceso de evaluación pone en evidencia múltiples aspectos, relacionados con las características y procesos de la institución educativa, los proyectos institucionales, los estilos de gestión, las propuestas curriculares y editoriales, las particularidades de los docentes y de los alumnos. La multidimensionalidad de los actos evaluativos, producto de sus articulaciones y sobre determinaciones, conduce a poner en duda la unidad y coherencia de lo que se designa con el término “evaluación”.

Acorde a los documentos elegidos, es menester resaltar la noción de evaluación que se presenta formalmente desde los distintos ministerios. Por un lado, en Finlandia, la metodología de evaluación ha abandonado las memorizaciones del sistema educativo y hace énfasis en el desarrollo de la curiosidad, la creatividad y la experimentación: no es una cuestión de transmitir información, sino que es más importante aprender a pensar (Finnish National Board of Education, 2003).

Por otro lado, el Gobierno de la CABA plantea que se evalúa de manera tal de construir un mapa del estado actual del progreso de los aprendizajes de los estudiantes para poder planificar de acuerdo al logro de los objetivos. Este tipo de evaluación permite que el docente cuente con la información necesaria acerca del progreso significativo de sus estudiantes (Ministerio de Educación de la Ciudad de Buenos Aires, 2013).

A continuación, se hará un análisis exhaustivo de los documentos a ser trabajados en función de distintas nociones que atravesarán y moldearán el artículo: la noción de “leer lo global” propuesta por Cowen (2000), la noción de “traveling libraries” (Popkewitz, 2005) y la noción de “making-up people” (Hacking, 2006).

4.2. Cowen: “leer lo global”

Un primer punto a señalar para desarrollar el análisis de una manera adecuada es la necesidad de “leer lo global” (Cowen, 2000). El proceso de conocer los discursos universales que perduran hoy en día en relación a la evaluación de los aprendizajes es indispensable para luego observar las dinámicas educacionales de cada región.

Un estudio comparativo que refiere a una temática particular “se concentra alrededor de estas ideas unificadoras y los académicos que leen lo global en diferentes momentos piensan en diferentes versiones de esto” (Cowen, 2005, p. 16). Por lo tanto, ¿cuáles son esos discursos externos, universales, globalizados, que juegan como punto de referencia en el campo de la evaluación de los aprendizajes? ¿Qué se dice sobre la evaluación de los aprendizajes?

Si bien al mirar en profundidad y analíticamente los documentos tiene especial preponderancia la evaluación formativa, se admite un discurso más amplio que se extiende por lo que plantean los documentos locales. Philippe Perrenoud (2010) afirma que “las pedagogías diferenciadas están ya a la orden del día y la evaluación formativa ya no es una quimera, pues ha dado lugar a numerosos ensayos en diversos sistemas” (p. 15). Es decir, parecería ser que la evaluación formativa, en la contemporaneidad de nuestros tiempos, permanece como “lo global”.

Esto se debe a que, por ejemplo, Boud y Falchikov (2005) plantean que “aunque la función formativa es a menudo eclipsada a diario por el predominio de las evaluaciones sumativas, se ha producido una renovación de intereses en la evaluación formativa en los últimos años”¹ (p. 35). Y más aún, diversos investigadores como Black y Wiliam (1998) o Elwood (2007) identifican en la evaluación formativa, lo que ellos denominan como “buenas prácticas”.

Entonces, considerando que el discurso de la evaluación formativa aparece en los documentos a analizar y se repite este patrón en la literatura contemporánea sobre evaluación, podríamos constatar que la evaluación formativa permanece como “lo global”, como lo que circula hoy en el mundo a través de las distintas regiones. Tanto los documentos como los textos utilizados definen, en mayor o menor medida, a la evaluación formativa como una actividad sistemática y continua que proporciona información necesaria sobre el proceso educativo y que orienta a los estudiantes a través de la retroalimentación.

El desafío está en la “reconstrucción en estos procesos sociales que suponen la destrucción del pasado y la redefinición del futuro” (Cowen, 2000, p. 35); es decir, este nuevo discurso admite la posibilidad de dejar a un lado las evaluaciones tradicionales a favor de la selección y construir evaluaciones modernas a favor del aprendizaje (Perrenoud, 2010).

A través de la lectura de lo global se puede identificar que los argumentos en torno a la evaluación en las diferentes regiones surgen gracias a la interdependencia mundial: acorde a cómo evolucione el mundo, acorde a las cuestiones que se presentan, acorde a las diferentes estrategias de evaluación que existan y perduren, acorde a los resultados académicos de las naciones, es que las concepciones locales en torno a la evaluación cambiarán o permanecerán para poder lograr mantener una educación actualizada y significativa en la formación integral del estudiante. En consecuencia, necesitamos de la lectura de lo global para comparar las tendencias de evaluación de ambas regiones analizadas y verificar la presencia o ausencia de este código educativo que predomina en el mundo actual.

Este nuevo patrón educativo sobre el que se trabaja -la evaluación formativa- puede percibirse de maneras diversas en los documentos con los que se busca trabajar. Por un lado, en el caso de Finlandia, a la hora de definir el currículum que va a regir para la escuela secundaria, se expone la necesidad de garantizarle a los estudiantes la oportunidad de completar sus estudios de manera flexible en función de su progreso y del feedback que le den sus guías académicos o docentes. Es decir, por detrás de los objetivos se esconde el propósito de que los alumnos tomen sus propias decisiones, si

¹ Traducción realizada por la misma autora. El texto original está en inglés.

dejar de aprovechar los aportes constructivos que le ofrezcan los docentes (Finnish National Board of Education, 2003)².

Los argumentos globales de la evaluación formativa están presentes en el documento de Finlandia porque se percibe al aprendizaje como el proceso a través del cual los estudiantes interpretan la información recibida en interacción con otros alumnos, con los profesores y con el entorno y en base a sus conocimientos previos ya incorporados (esta es la noción de aprendizaje significativo). A pesar de que se promueva el trabajo independiente, los alumnos tienen que ser guiados para ser conscientes de su propio proceso de enseñanza-aprendizaje.

La educación en Finlandia tiene en cuenta el hecho de que las habilidades de los estudiantes varían y por eso necesitan de los docentes para que supervisen sus estudios de maneras diferentes. La presencia de cierto grado de heterogeneidad nos permite afirmar que Finlandia hace hincapié en los procesos individuales de cada alumno, teniendo en cuenta la importancia de los procesos formativos para crear, a su vez, "*life-long learning and continuous self-development*" (Finnish National Board of Education, 2003, p. 12).

En conclusión, desde el Ministerio de Educación de Finlandia, se promueve el discurso global de evaluación formativa ya que definen a la evaluación como el proceso a través del cual se promueve la auto-evaluación en complemento con el feedback formativo para identificar el progreso y los resultados de aprendizaje: "the role of assessment is to provide students with feedback on how well they have met the objectives of the course and on their progress in that subject" (Finnish National Board of Education, 2003, p. 224).

Por otro lado, desde el Ministerio de Educación de la CABA, también pueden identificarse argumentos que se inscriben en el paradigma de la evaluación formativa sobre todo por la insistencia en generar aprendizaje significativo. Además, al igual que en Finlandia, se propone que el sistema educativo adquiera la flexibilidad suficiente como para asumir escenarios cambiantes y poder brindarle a los estudiantes la habilidad de aprender durante toda la vida. Las nociones de *life-long learning* y de aprendizaje significativo son indispensables en el marco de la evaluación formativa.

El sistema educativo de la CABA admite entonces que la evaluación debe ser "formativa, formadora y sumativa" (Ministerio de Educación de la Ciudad Autónoma de Buenos Aires, 2013, p. 61) de manera tal de poder comprender los procesos y los resultados para mejorar los procesos de aprendizaje de los alumnos. Es decir, en el documento formal -el Diseño Curricular- se presenta el discurso global de evaluación formativa ya que se promueve el deseo de proporcionar una mirada integral y de mejorar los procesos de enseñanza a través de la devolución fundamentada de docentes y de los propios compañeros de curso (p. 145). También, como parte del discurso de la formatividad, se explica cómo este tipo de evaluación permite conocer lo que el estudiante sabe hacer y qué es lo que le falta aprender, es decir, conocer su progreso y motivarse con sus logros.

Si bien desde la Ciudad de Buenos Aires el discurso no se orienta, como en Finlandia, a la individualización de las trayectorias escolares, se admiten procesos heterogéneos

² Cita textual: "flexible arrangements for progression of studies and provision of special support as required. The curriculum must be drawn up so as to provide students with an opportunity to make individual choices" (2003, p. 8).

orientados al seguimiento de los aprendizajes de los estudiantes con el objetivo de perfeccionar las trayectorias, “brindando una mayor autonomía y responsabilidad a los estudiantes” (Ministerio de Educación de la Ciudad Autónoma de Buenos Aires, 2013, p. 146). De esta manera, el trabajo se orienta, en el marco de la evaluación formativa, a trabajar con los errores de los alumnos de manera constructiva para generar aprendizaje significativo.

4.3. Popkewitz: “Traveling libraries”

¿Cómo se realiza la traducción de este discurso global? “¿Cómo se realiza la traducción de aquello que es externo al colegio?” (...) “¿Cómo pasa de ser una idea externa a ser una práctica educativa?” (Cowen, 2005, p. 17).

Para que la evaluación formativa se ponga en juego, no es suficiente conocer de qué se trata y plantearlo como un paradigma teórico a tener en cuenta. La evaluación formativa se pone en juego en el momento en el que ocurre la traducción, en el momento en el que se conjugan elementos externos e internos.

Para combinar el discurso teórico con lo que sucede en el interior de los sistemas educativos escogidos, es necesario tener en cuenta el concepto de “traveling libraries” (Popkewitz, 2005). Este autor demuestra que el viaje de los discursos no es desde un centro a la periferia ni de un punto fijo a otro punto fijo, sino más bien, está mediado por los marcos discursivos socio-históricos y específicos que redefinen el interés de cada nación. Esto permite comprender que la tendencia visible y popular de la evaluación formativa que se traslada actualmente alrededor del mundo, es reinterpretada y redefinida por cada región a su manera.

La noción de “travelling libraries” da especial importancia al entramado de redes a través del cual se da inteligibilidad a los cambios y proporciona una manera de pensar acerca de lo que aparece en la superficie de los debates educativos contemporáneos. Es decir, este concepto proporciona una manera de considerar la globalización y la localización como procesos relacionados y superpuestos de manera diferente en los distintos lugares.

En función de estos aportes y de lo presentado como evaluación formativa desde los diferentes documentos, podemos analizar las dimensiones socio-históricas que delimitan el campo de la evaluación en los sistemas educativos. Es decir, la manera en la que “aterriza” el discurso de la evaluación en las distintas regiones estará influenciado por los intereses que tenga cada nación.

Por un lado, en el caso de la Ciudad de Buenos Aires, a lo largo de los años, la escuela media se ha enfrentado a numerosos cambios en el campo pedagógico que hacen que, en la actualidad, tenga un significado distinto al que tenía en el momento fundacional de los sistemas educativos occidentales. Tras la promulgación de la Ley Nacional de Educación 26.206 del 2006, el nivel medio o secundario se ha convertido en el nuevo piso de la escolaridad obligatoria (Tenti, 2000, p. 2). Esta nueva disposición afecta notablemente los modos de hacer las cosas en el interior de las instituciones, dando lugar a una serie de transformaciones en los dispositivos y procesos escolares.

El enfrentamiento a una nueva disposición de los contenidos escolares, de las estrategias de enseñanza, de los actores participantes dentro del sistema, exige a los alumnos incorporar ciertas habilidades que se correspondan con los lineamientos del siglo XXI: “necesidad de acercar la situación de la escuela secundaria a la realidad

compleja del siglo XXI, cuya sociedad es más flexible y cambiante” (Ministerio de Educación de la Ciudad Autónoma de Buenos Aires, 2013, p. 73).

La obligatoriedad de la escuela secundaria, entonces, introduce la necesidad de transformar el proceso de enseñanza-aprendizaje, buscando promover que el aprendizaje en este nivel sea significativo de manera tal que se construyan alumnos autónomos y capaces de razonar de manera personal:

El Ministerio de Educación de la Ciudad Autónoma de Buenos Aires ha puesto el eje de toda política educativa en el aprendizaje significativo de los estudiantes. Se ha iniciado un proceso de llevar el sistema educativo orientado en el docente a un sistema orientado en el aprendizaje del alumno y la nueva organización institucional necesaria para que el aprendizaje suceda. (Ministerio de Educación de la Ciudad Autónoma de Buenos Aires, 2013, p. 25)

Para promover y contribuir al aprendizaje significativo, resulta necesario hacer hincapié en el campo de la evaluación ya que, para elevar la calidad de los aprendizajes, el rol de la evaluación educativa es central: es un componente esencial del proceso que transcurre en la escuela.

En la contemporaneidad de nuestros tiempos, la evaluación se presenta como una problemática educativa que, a nivel micro, forma parte del debate didáctico y constituye uno de los problemas centrales para los docentes de la escuela media. Con el objetivo de transformar la evaluación tradicional, los docentes intentan visibilizar las adquisiciones y los modos de razonar de cada alumno haciendo de la enseñanza, una más personalizada.

La evaluación está en proceso de dejar a un lado su función selectiva y aumentar una función pedagógica. A pesar de que todos los tipos de evaluación sean indispensables para la mejora en la calidad de los aprendizajes, se reconoce a la evaluación formativa como la práctica crucial para generar aprendizaje significativo: “la evaluación de carácter formativo permite al estudiante conocer qué es aquello que sabe hacer y qué es lo que le falta aprender (...) debe estar orientada al seguimiento de los aprendizajes de los estudiantes con el objetivo de perfeccionar las trayectorias escolares” (Ministerio de Educación de la Ciudad Autónoma de Buenos Aires, 2013, p. 146).

Sin embargo, tanto los docentes como los directivos se enfrentan a diario con dificultades que se presentan a la hora de definir los criterios de evaluación. Frente a las adversidades de la actual escuela secundaria argentina, este tipo de evaluación está lejos de ser puesta en práctica en todas las instituciones con coherencia, ya que lucha contra los numerosos obstáculos institucionales: el número de alumnos, la sobrecarga de los programas, la concepción de los medios de enseñanza y las didácticas.

Por otro lado, tomando los aportes de Simola (2005), podemos comprender el marco socio-histórico en el cual se inscribe la evaluación formativa en Finlandia. En las últimas décadas, Finlandia ha desarrollado políticas educativas que lo llevaron en pos del éxito y del reconocimiento internacional.

El alto rendimiento del sistema finlandés parece ser atribuible a toda una red de factores relacionados entre sí: las actividades de interés y de ocio, las oportunidades de

aprendizaje proporcionadas por las escuelas, el apoyo de los padres y la participación³, el contexto social y cultural del aprendizaje y de los estudiantes de todo el sistema educativo. Pero la historia de éxito de Finlandia es sumamente reciente.

El sistema de enseñanza secundaria se desarrolló sólo en la década de los 70', pero al mismo tiempo se llevó a cabo muy rápidamente y de forma sistemática, incluso de una manera bastante totalitaria. La enseñanza y el aprendizaje siempre fueron muy tradicionales e implicaron la exposición por parte del docente en frente de todo el grupo de alumnos (Simola, 2005). A pesar de que la cultura finlandesa aún tenga incorporado un elemento significativo de mentalidad autoritaria, conservadora y obediente, hoy en día las cosas han llegado a ser más independientes y, en consecuencia, la evaluación sigue una pedagogía más independiente: "each student will draw up his or her own individual study plan" (Finnish National Board of Education, 2003, p.8) y, además, "assessment will encourage students in a positive way to set their own objectives and to readjust their working methods" (p. 224).

La cultura arcaica prevalece pero hay una presencia de confianza social y de apreciación a los maestros que cumplen su labor de manera comprometida. En la actualidad, las dinámicas en el aula están organizadas en función de trabajo individual o grupal de los alumnos mientras que el docente los guía dándole feedback al trabajo que van realizando. Por lo tanto, en este nuevo contexto donde se pone en juego no sólo la interacción docente-alumno sino también la de alumno-alumno, la evaluación formativa toma un rol indispensable.

La evaluación alentará a los estudiantes de una manera positiva para establecer sus propios objetivos y reajustar sus métodos de trabajo. A su vez, proporcionará a los estudiantes con información sobre lo bien que se han cumplido los objetivos del curso y sobre los progresos realizados en cada una de las materias.

A diferencia de la CABA, en Finlandia la evaluación se inscribe en un contexto de confianza, donde el control y la regulación tiene un rol poco observable. Esto se debe más que nada al desarrollo de las políticas educativas que se dieron a partir de 1995. Desde principios de los 90', el discurso educativo estatal se ha centrado en la evaluación como la herramienta fundamental para un desarrollo de calidad. Mientras que anteriormente se había creído que las metas educativas podían alcanzarse ajustándose a normas estrictas, la convicción en los 90' fue que su consecución requería el establecimiento de objetivos nacionales básicos y la evaluación de los logros a la luz de los posteriores resultados.

En conclusión: "the Finnish 'secret' of top-ranking may therefore be seen as the curious contingency of traditional and post-traditional tendencies in the context of the modern welfare state and its comprehensive schooling" (Simola, 2005, p. 466).

4.4. Hacking: "making-up people"

Ahora bien, la interacción entre los discursos cosmopolitas democráticos y las reformas educativas actuales o contextos locales (Friedrich, 2014) admite un proceso de traducción y recontextualización de lo que se conoce como evaluación formativa (Ball, 2002). Fue de especial relevancia atender a las particularidades locales y a la

³ "Students' parents or guardians must be provided with sufficient opportunities to familiarise themselves with the work of the upper secondary school. Co-operation with parents will support the prerequisites for study" (Finnish National Board of Education, 2003, p. 18).

formulación y promulgación de políticas para identificar cuáles son las similitudes o diferencias con los patrones generales y visibilizar la manera en la que se pone en juego la noción de “*traveling libraries*”.

No obstante, sabiendo que por detrás de cualquier acto evaluativo existe un objetivo por parte del docente que lleva a cabo la acción, queda por explorar el tipo de sujeto con el que se trabaja en las distintas regiones. Como respaldo teórico, tomaremos la noción de “*making-up people*” (Hacking, 2006) para enriquecer el argumento de que, con la evaluación formativa, se construyen nuevos participantes antes inexistentes.

Por un lado, en lo que respecta al Diseño Curricular de Finlandia, se ha destacado la presencia indispensable de la confianza y la libertad. La evaluación en esta región no sólo admite ciertas ideas y características de la evaluación formativa, sino que presenta un alumno libre y capaz de regular sus propios aprendizajes de manera autónoma, con la seguridad de tener una guía constructiva que lo ayude a alcanzar sus objetivos de aprendizaje: “*students are seen to be the constructors of their own learning, competence and views of the world*” (Finnish National Board of Education, 2003, p.12).

En consecuencia, el docente tiene una postura que, en palabras de Fenstermacher y Soltis (1998), podría denominarse “*terapeuta*”: un docente cuyo énfasis está puesto en el alumno y en lo que él puede lograr, en su desarrollo psíquico e intelectual para escoger la ruta que más potencie el aprendizaje del estudiante: “*counselling and guidance from an integrated whole, which aims to support students at different stages of their upper secondary school studies and to develop their capabilities to make choices and find solutions concerning education, training and future career*” (Finnish National Board of Education, 2003, p.18).

Por otro lado, en el caso de la CABA, el docente se asocia al paradigma del “*ejecutivo*” (Fenstermacher y Soltis, 1998): un docente que planifica, que recibe instrucciones, que lo orientan a la hora de poner en juego ciertas estrategias y cuyo rol es fundamental para que la clase tenga sentido y ocurra. Existe un documento de carácter orientativo para los docentes que “*propone un desarrollo más exhaustivo que incluye los componentes del material anterior -propósitos y contenidos- junto con la presentación de los objetivos de aprendizaje por año de estudio y alcances y comentarios para el desarrollo de los contenidos, incluyendo sugerencias para su enseñanza*” (Ministerio de Educación de la Ciudad Autónoma de Buenos Aires, 2013, p. 36).

Por su parte, el alumno posee menos libertad a la hora de elegir sus propios cursos para completar la escolaridad secundaria obligatoria y, además, está sumamente influenciado por las aptitudes que, desde el Ministerio, esperan que tenga el futuro egresado de la NES: “*metas que todos los egresados del nivel secundario DEBEN alcanzar. Dichas aptitudes pueden definirse como los recursos cognitivos y comportamentales que debería adquirir el alumno en el trayecto de su educación media para enfrentar las demandas de los estudios superiores*” (Ministerio de Educación de la Ciudad Autónoma de Buenos Aires, 20013, p. 74).

En resumen, si vivimos en un mundo en donde los objetivos se mueven constantemente, la interacción entre un docente y un alumno -por ejemplo- se va a ver modificada. Y ya que cambian las interacciones y los objetivos de la evaluación, las personas ya no son exactamente la misma clase de gente que eran antes.

A pesar de Finlandia y Buenos Aires tener objetivos distintos en cuanto a la educación, las estrategias de enseñanza y los aspectos más delicados de la evaluación, en los

documentos ambos dos poseen un capítulo específico que refiere a la evaluación y otro capítulo que refiere a las habilidades o aptitudes a lograr por parte del docente y el alumno. Esa similitud podría considerarse como un aspecto que hoy está globalizado en los diseños curriculares alrededor del mundo.

5. Conclusión

A lo largo de este trabajo académico se ha intentado exponer, describir y explicar las distintas concepciones construidas en torno a la temática de la evaluación. Siguiendo los argumentos propuestos por ciertos investigadores de educación comparada, se ha hecho un estudio comparativo entre Finlandia y la Ciudad Autónoma de Buenos Aires, que permite visibilizar los modos y las configuraciones que se inscriben en la noción de evaluación formativa de cada región.

La propuesta de este artículo ha sido complejizar las ideas formales de evaluación que se presentan en los Diseños Curriculares de las distintas regiones. De tal manera, se intentó percibir cómo “aterizan” las características de la evaluación formativa en Finlandia y Buenos Aires, cómo se traslada el concepto de evaluación formativa al campo formal de los Diseños Curriculares.

El objetivo primordial fue identificar cómo se leen las propuestas o los argumentos de evaluación, sabiendo que los docentes que están al servicio de la educación y la escuela poseen sus propias estrategias de enseñanza y cargan con un tinte de subjetividad. Es decir, explicar cómo se configuran los enfoques, cuáles son las similitudes y diferencias entre las regiones y qué orientaciones se da a los profesores a la hora de poner en juego la evaluación.

Una primera aproximación refleja que la evaluación existe desde que se creó el dispositivo escuela y, en tanto estrategia global, se hacen distintas lecturas de la misma en las diferentes regiones analizadas. De todas maneras, ambos inscriben la noción de evaluación formativa en sus estrategias de evaluación de los aprendizajes y hacen hincapié en la individualización de los procesos escolares acorde a las capacidades de los alumnos. Además, intentan construir individuos que aprendan a lo largo de toda la vida, es decir, *life-long learners*.

También ambos explican el deseo de que los alumnos sean individuos autónomos, capaces de incorporarse al mundo laboral sin dificultades sino, más bien, con las capacidades suficientes como para enfrentar lo cotidiano de la vida real. Sin descartar que las tareas escolares están en constante evolución -por los avances tecnológicos y las mejores en el campo pedagógico- el dispositivo de evaluación sigue existiendo en todas las escuelas a lo largo del mundo, de alguna u otra manera.

Ahora bien, ¿qué se puede rescatar de Finlandia que no se haya puesto en juego aún en la Ciudad de Buenos Aires? El argumento de la evaluación formativa ya no es algo para o de algunos pocos sino que está, cada vez más, girando alrededor del mundo. La diferencia reside, como fue explicado anteriormente, en el grado de utilización de las características o en la manera en la que la ponen en juego tanto los docentes como los alumnos.

Entonces, afirmando que la evaluación formativa aumenta la calidad de los aprendizajes, la CABA debería formar a los docentes de manera más terapéutica y no tan ejecutiva. Es así como el docente incorporará el ejercicio de reflexionar sobre el

aprendizaje de sus alumnos –como actitud moderna– para luego guiarlos en la construcción de su autonomía.

Referencias

- Acevedo, P. A. (2001). *La evaluación en una concepción de aprendizaje significativo*. Valparaíso: Ediciones Universitarias de Valparaíso – Pontificia Universidad Católica de Valparaíso.
- Apel, J. (1998). *Evaluar e informar en el proceso de enseñanza-aprendizaje*. Buenos Aires: Aique Grupo Editor.
- Ball, S. (2002). Grandes políticas, un mundo pequeño. Introducción a una perspectiva internacional en políticas educativas. En M. Narodowski, M. Nores, y M. Andrada (Coord.), *Nuevas tendencias en Políticas Educativas* (pp. 103- 129). Buenos Aires. Ediciones Juan Granica.
- Black, P. y Wiliam, D. (1998). *Inside the black box: Raising standards through classroom assessment*. Londres: Kings College.
- Boud, D. y Falchikov, N. (2005). Redesigning assessment for learning beyond higher education. *Higher Education Research and Development Society of Australasia*, 28, 34-41.
- Cowen, R. (2005). Los códigos secretos de los sistemas educativos: Leyendo las “Piedras Roseta”. *Boletín de la Academia Nacional de Educación*, 63, 12-20.
- Elwood, J. (2007). Formative assessment: Possibilities, boundaries and limitations. *Assessment in Education: Principles, Policy & Practice*, 13(2), 215-232. doi:10.1080/09695940600708653.
- Finnish National Board of Education. (2003). *National Core Curriculum for Upper Secondary Schools*. Helsinki: Studio Viiva Oy.
- Fenstermacher, G. y Soltis, J. (2001). *Enfoques de la enseñanza*. Buenos Aires: Amorrortu.
- Friedrich, D. (2014). Global microlending in education reform: Enseñá por Argentina and the neoliberalization of the grassroots. *Comparative Education Review*, 58(2), 296-321. doi:10.1086/675412.
- Foucault, M. (1984). Qu'est-ce que les Lumières? En J. Dávila (Ed.), *¿Qué es la Ilustración?* (pp. 19-46). Madrid: Actual.
- Giddens, A. (1990). *Consecuencias de la modernidad*. Madrid: Alianza Editorial.
- Hacking, I. (2006). Making up people. *London Review of Books*, 28(16), 23-26.
- Kant, I. (1784). Respuesta a la pregunta: ¿qué es la Ilustración? *Revista Colombiana de Psicología*, 14, 7-10.
- Meyer, J. y Ramírez, F. (2000). La institucionalización mundial de la educación. En J. Schriewer (Comp.), *Formación del discurso en la educación comparada* (pp. 91-111). Barcelona: Pomares.
- Ministerio de Educación de la Ciudad Autónoma de Buenos Aires. (2013). *Nueva Escuela Secundaria de la Ciudad de Buenos Aires. Diseño Curricular. Ciclo Básico 2014-2020*. Recuperado de http://www.buenosaires.gob.ar/areas/educacion/recursos/NESCB-2014_web.pdf2013
- Perrenoud, P. (2008). *La evaluación de los alumnos. De la producción de la excelencia a la regulación de los aprendizajes*. Buenos Aires: Colihue.
- Pineau, P., Dussel, I. y Carusso, M. (2001). *La escuela como máquina de educar. Tres escritos sobre un proyecto de la modernidad*. Buenos Aires: Paidós.

- Popkewitz, T. (2013). Pisa, números, estandarización de la conducta y la alquimia de las materias escolares. *Profesorado*, 17(2), 47-64.
- Scriven, M. (1967). The methodology of evaluation. En Tyler, R. W., Gagné, R. M. y Scriven, M. (Eds.), *Perspectives of curriculum evaluation*, 39-83. Chicago, IL: Rand McNally.
- Simola, H. (2005). The Finnish miracle of PISA: Historical and sociological remarks on teaching and teacher education. *Comparative Education Review*, 41(4), 455-470. doi:10.1080/03050060500317810.
- Tedesco, J. C. y López, N. (2002). Desafíos a la educación secundaria en América Latina. *Revista CEPAL*, 76, 55-69.
- Tenti, E. (2000). *Culturas juveniles y cultura escolar*. Buenos Aires: IIPE-UNESCO.

Concepciones Sobre Evaluación en la Formación Inicial del Profesorado de Ciencias, Tecnología y Matemáticas

Assessment's Conceptions of Pre-service Secondary Science, Technology and Maths Teachers

Alfonso Pontes Pedrajas^{1*}, Francisco J. Poyato López¹ y José M^a. Oliva Martínez²

⁽¹⁾ Universidad de Córdoba y ⁽²⁾ Universidad de Cádiz

En este trabajo se muestran los resultados de un proyecto de innovación docente sobre evaluación educativa, desarrollado en el Máster de formación del profesorado de educación secundaria, aplicando una metodología de enseñanza basada en el enfoque reflexivo. Han participado, durante dos cursos académicos, 71 estudiantes de varias especialidades del área científico-técnica, realizando en el aula actividades que han permitido explicitar y debatir sus ideas previas sobre la evaluación del aprendizaje. Mediante un análisis de tipo cualitativo se han estudiado las opiniones de los participantes en torno a un conjunto de cuestiones abiertas, lo cual nos ha permitido conocer las concepciones de los futuros docentes sobre la evaluación, analizar su relación con diversos modelos didácticos y valorar el grado de consistencia de tales ideas. La experiencia ha resultado positiva porque nos ha permitido conocer mejor el pensamiento inicial de los participantes sobre la evaluación educativa. También nos sirve de estímulo para seguir trabajando en un proceso de formación en el que los futuros docentes realizan actividades orientadas a reflexionar sobre problemas de interés educativo, explicitando sus ideas previas y debatiendo sobre las diferentes concepciones como condición necesaria para impulsar la progresión del pensamiento profesional docente.

Palabras clave: Profesorado de enseñanza secundaria, Formación inicial, Concepciones previas, Evaluación educativa, Educación científico-técnica.

In this work, we show the results of a project of teaching innovation on educational evaluation, developed at the Master of Teacher Training Secondary Education, applying a methodology of teaching based on the reflexive approach. In this project, 71 students from various specialties of the scientific-technical area have participated during two academic courses. They have carried out activities that allow us to specify and discuss their previous ideas on assessment of learning. We have studied the opinions of the participants through a qualitative analysis of a set of open questions. This has enabled us to understand the conceptions of future teachers about the evaluation and, in addition, to analyze its relationship with various teaching models and to assess the degree of consistency of such ideas. The experience has been positive because it has permitted us to get a better understanding of the initial thinking of participants about the educational evaluation. It also encourage us to continue with this work, in which the future teachers put in practice activities aimed to reflect on issues of educational interest, describe their previous ideas and debate about the different conceptions, as a necessary condition to promote the progression of teacher professional thinking.

Keywords: Secondary education teachers, Initial training, Previous conceptions, Evaluation, Scientific and technical education.

1. Introducción¹

El desarrollo y puesta en marcha del llamado Espacio Europeo de Educación Superior (EEES) ha supuesto para la Universidad Española un proceso de transformación y convergencia de la enseñanza superior, que también ha afectado al curso de Formación Inicial del Profesorado de Secundaria (FIPS), ya que el anterior modelo de formación inicial, basado en el curso para la obtención del Certificado de Aptitud Pedagógica (CAP) presentaba numerosas limitaciones y había quedado obsoleto (Perales, 2014). Desde el curso 2009-10 la FIPS se ha integrado de forma oficial en los estudios universitarios de postgrado y se ha transformado en el Máster de Profesorado de Enseñanza Secundaria (MPES), donde se pretenden desarrollar competencias docentes de carácter profesional en un marco formativo más amplio y mejor organizado. Este hecho ha supuesto un importante cambio en la estructura del plan de estudios de la FIPS y en los programas de las materias de los diversos módulos teóricos y de la fase práctica, pero también debería aprovecharse para renovar los métodos de formación e impulsar la innovación y la investigación educativa en el campo de la formación docente (García-Carmona, 2013; Serrano y Pontes, 2015).

En este contexto han cobrado auge los trabajos orientados a mejorar el conocimiento de los profesores de secundaria en formación inicial, investigando sus motivaciones, actitudes, expectativas y, sobre todo, sus concepciones sobre los múltiples aspectos que configuran la profesión docente (e.g. Brown, Lake y Matters, 2011; Meirink, Meijer, Verloop y Bergen, 2009), con objeto de diseñar propuestas metodológicas que favorezcan el cambio de las ideas previas de los futuros docentes y les ayuden a construir conocimientos profesionales adecuados para desarrollar la actividad docente en los actuales centros de educación secundaria (Porlán y Martín, 2006; Sethusha, 2013). La finalidad principal de este trabajo consiste en plantear actividades que permitan a los futuros docentes de enseñanza secundaria reflexionar sobre la problemática de la evaluación del aprendizaje (De Wries, 2015; Turpo-Gebera, 2011; Vergara, 2012), permitiendo explicitar sus concepciones previas sobre este tema y favorecer el cambio conceptual (Abell, 2007; Hewson, 1993). Al mismo tiempo pretendemos recoger datos que nos permitan conocer las características de tales concepciones, por considerar que pueden ser útiles para el avance de la investigación sobre el pensamiento del profesorado (Carbajal y Gómez, 2002; Hernández, 2010).

2. Fundamento y antecedentes

2.1. Modelos didácticos y concepciones del profesorado sobre los procesos educativos

La práctica docente está influenciada en gran medida por el pensamiento de los profesores sobre los procesos educativos y para mejorar la calidad de la educación es necesario mejorar el proceso de formación del profesorado, teniendo en cuenta -entre

¹ Ayudas: Este trabajo forma parte de un proyecto de innovación y mejora de la calidad docente, desarrollado en sucesivas fases, contado con la financiación del Vicerrectorado de Planificación y Calidad Educativa de la Universidad de Córdoba (España).

otros aspectos- el papel que desempeñan las concepciones personales en los procesos de formación (Abell, 2007; Mellado, Blanco y Ruiz, 1999; Porlán, Rivero y Martín del Pozo, 1997). Se han realizado muchas investigaciones orientadas a detectar las concepciones previas y creencias del profesorado, en formación inicial o en activo, utilizando diferentes tipos de instrumentos (Contreras, 2010; Valbuena, 2007) y en relación con los antecedentes de nuestra investigación presentan especial importancia los trabajos sobre el pensamiento docente que abordan problemas educativos de interés para los futuros docentes, desde un enfoque reflexivo (Fuentes, García y Martínez, 2009; García-Carmona, 2013; Pontes, Poyato y Oliva, 2015). También nos parecen interesantes aquellos estudios en los que se detectan dificultades de aprendizaje o los que permiten representar el cambio producido en las concepciones docentes durante la formación inicial (Solís, Porlán y Rivero, 2012). Consideramos que los formadores del profesorado pueden utilizar los resultados de tales investigaciones para promover la progresión de las ideas previas y ayudar a construir un conocimiento profesional adecuado (Dixon y Haigh, 2009).

En varios de los trabajos precedentes se han logrado identificar algunos perfiles característicos del pensamiento y la acción docente, que pueden interpretarse en términos de modelos didácticos diferenciados, ya que ofrecen una visión propia y distintiva de los procesos educativos en el área científico-técnica (Fuentes et al., 2009; Porlán et al., 1997; Pontes et al., 2015; Solís et al., 2012). Existe mucha literatura sobre el tema de los modelos didácticos y su relación con el pensamiento docente, pero hay trabajos de revisión bibliográfica que ofrecen una panorámica global e interesante sobre el tema (Hernández y Maquilón, 2010). En un trabajo de síntesis sobre los diferentes perfiles del profesorado de ciencias, llevado a cabo por Oliva (2008), se han caracterizado cuatro modelos didácticos que responden a las denominaciones genéricas siguientes: (1) Tradicional, (2) Tecnológico, (3) Activista y (4) Socio-constructivista. Los modelos 1 y 4 están muy diferenciados, pues representan visiones contrapuestas del pensamiento docente, ya que el modelo tradicional se identifica con un enfoque educativo basado en la transmisión-recepción mientras que el modelo socio-constructivista presenta un carácter innovador y se relaciona con un proceso de enseñanza-aprendizaje basado en la construcción de nuevas ideas a partir de los conocimientos previos. Sin embargo, los modelos 2 y 3 pueden considerarse como modelos intermedios o híbridos, ya que a veces incluyen características próximas a los modelos situados en los extremos de la escala anterior. Los estudios que permiten relacionar las concepciones previas del profesorado con los diferentes modelos didácticos que caracterizan el pensamiento docente nos parecen interesantes porque tales modelos permiten visualizar de forma más clara la progresión del conocimiento que se produce a través del proceso de formación (Fuentes et al., 2009, Valbuena, 2007).

En la línea de investigación sobre el pensamiento docente, hemos realizado anteriormente varios estudios sobre las concepciones de profesores en formación inicial, en torno a diversas estrategias y recursos educativos relacionados con el enfoque reflexivo (Pontes, Serrano, Muñoz y López, 2011) o la profesionalidad docente (Pontes, Serrano y Poyato, 2013). En un trabajo reciente hemos explorado las concepciones de profesores en formación sobre el aprendizaje de las ciencias (Pontes et al., 2015) y hemos encontrado que tales concepciones se agrupan mayoritariamente en torno a dos enfoques educativos diferenciados que hemos denominado “tradicional” e “innovador”, aunque también hay muchas ideas de carácter dual o ambivalente que podrían formar parte de un enfoque intermedio. El enfoque innovador se aproxima

bastante al modelo socio-constructivista, aunque los profesores en formación no presentan un esquema de ideas tan estructurado y coherente como el modelo didáctico representado bajo esta denominación (Contreras, 2010; Oliva, 2008), de modo que hemos optado por usar un término menos comprometido, aunque somos conscientes de que el enfoque innovador puede resultar un tanto ambiguo. En este nuevo estudio vamos a seguir avanzando en esta línea centrándonos en explorar las concepciones previas de los estudiantes del MPES sobre la evaluación del aprendizaje.

2.2. Concepciones del profesorado acerca de la evaluación del aprendizaje

De Vries (2015) ha analizado recientemente las múltiples facetas de la evaluación educativa y ha destacado la importancia del proceso de evaluación del aprendizaje para mejorar la calidad de la enseñanza, indicando que este tema debe ser tratado con cuidado por su complejidad siendo interesante impulsar la investigación en este campo. Uno de los aspectos a considerar, a la hora de mejorar el proceso de evaluación del aprendizaje, es el papel que desempeñan las concepciones, creencias y teorías implícitas del profesorado respecto de dicho proceso (Prieto y Contreras, 2008; Borko, 1997).

Desde hace tiempo se han realizado interesantes investigaciones sobre las concepciones del profesorado, en formación o en ejercicio, respecto a la evaluación del aprendizaje. Un trabajo pionero en España fue realizado por Buendía, Carmona, González y López (1999), quienes abordaron el estudio de las concepciones del profesorado de secundaria sobre la evaluación utilizando una batería de cuestiones abiertas. Tales autores ofrecen una síntesis interesante de las ideas de los docentes sobre la evaluación, que podríamos considerar bastante próximas al enfoque educativo tradicional porque hacen hincapié en términos tales como valorar, medir, sancionar o promocionar al alumnado, pero los profesores son conscientes de la importancia de la evaluación para mejorar los procesos de enseñanza-aprendizaje y algunos docentes también consideran que la evaluación es importante para hacer un seguimiento del proceso de construcción del conocimiento de los alumnos, lo cual se relaciona con un enfoque constructivista o innovador de la evaluación.

Otros autores como Segers y Tillema (2011) realizaron un estudio destinado a explorar las concepciones de profesores y alumnos de secundaria sobre la evaluación. Observaron que los alumnos no distinguen entre evaluación sumativa y formativa, pero los profesores sí hacen esa distinción. Los profesores tienen una visión más compleja pues consideran, entre otros aspectos, la responsabilidad de la escuela y del profesorado para mejorar la evaluación y, sobre todo, los efectos que la evaluación tiene en el proceso de aprendizaje. Por su parte, Remesal (2011) también ha realizado un estudio sobre las concepciones de profesores de secundaria y maestros acerca de la evaluación, elaborando un modelo de análisis de tales concepciones en el que se consideran cuatro dimensiones. Tales dimensiones permiten estudiar los efectos de la evaluación sobre la enseñanza, el aprendizaje, la responsabilidad del profesorado o de la escuela y la certificación de los resultados académicos alcanzados. El estudio puso de manifiesto la dificultad de poner en práctica un sistema de evaluación innovador y alternativo que ayude a mejorar el proceso de aprendizaje al profesorado de ambos niveles que participaba en el proyecto.

Monteagudo, Molina y Miralles (2015) han explorado recientemente las opiniones sobre la evaluación del profesorado español del área de ciencias sociales en la educación secundaria. Tales autores observan que los docentes en ejercicio muestran escasa formación en el ámbito evaluativo, y que continúan evaluando a los alumnos

básicamente mediante procedimientos tradicionales, como los exámenes escritos que enfatizan contenidos conceptuales. Según los citados autores el enfoque de competencias básicas está poco presente en las concepciones del profesorado y concluyen que las reformas legales educativas tienen escasa repercusión en la práctica evaluadora cotidiana de los profesores de este nivel.

En relación al tema de las concepciones del profesorado Dixon y Haigh (2009) realizaron un estudio cualitativo donde se observa la influencia mediadora de las creencias de un conjunto reducido de profesores de matemáticas en torno al proceso de evaluación y se analizan los cambios en el pensamiento docente sobre los roles de profesores y alumnos en relación a dicho proceso. Por su parte, Wang, Kao y Lin (2010) describen y analizan las concepciones sobre el aprendizaje y la evaluación de la ciencia. Los resultados muestran que el modo de evaluar de los profesores en formación estaba más próximo a una visión tradicional del aprendizaje.

En un contexto más cercano a nuestra investigación, Hernández (2010) ha realizado un estudio de las concepciones sobre evaluación de profesores de ciencias de secundaria en formación inicial, en el curso del CAP antes citado, utilizando una batería de cuestiones abiertas donde se abordan tópicos clásicos sobre el tema ya analizados por Buendía et al. (1999), aunque en este caso los participantes tenían que responder desde dos puntos de vista: según su experiencia real como alumnos y según su idea de cómo debería ser el proceso de evaluación. Entre las implicaciones educativas de este trabajo se destaca la necesidad de ayudar a los profesores en formación a reflexionar sobre la evaluación del aprendizaje desde una perspectiva formativa y motivadora. Por su parte, Turpo-Gebera (2011) ha analizado las concepciones y prácticas evaluativas de los docentes de secundaria del área de ciencias, usando un cuestionario como instrumento de aproximación al pensamiento docente. El autor indica que, a pesar de los procesos de formación y capacitación docente, subsisten intervenciones tradicionales en las formas y modos de evaluar los aprendizajes. En este estudio también se observa que, a través de las cuestiones planteadas, afloran posicionamientos relativamente innovadores sobre la evaluación, aunque en conjunto revelan situaciones de incertidumbre y ambivalencia.

En el ámbito anglosajón, Brown et al. (2011) han investigado las diferentes creencias sobre la evaluación del aprendizaje de las ciencias, entre profesores de enseñanza secundaria y primaria, poniendo de manifiesto la importante relación entre las creencias de los docentes y las diferentes prácticas de evaluación en ambos niveles de enseñanza. Los resultados mostraron que los maestros de primaria están más de acuerdo que los profesores de secundaria con la idea de que el sistema de evaluación usado por cada profesor determina la calidad del proceso de enseñanza y la mejora del aprendizaje. Así mismo Halinen, Ruohoniemi, Katajavuori y Virtanen (2014) han estudiado las concepciones sobre la evaluación en profesores de biología y han identificado tres categorías o modelos docentes atendiendo al discurso de cada profesor sobre el tema. Tales autores consideran que la formación pedagógica, la relación con los compañeros, el apoyo institucional y la interacción con la comunidad educativa influyen de diferente manera en el sistema de evaluación que cada docente realiza en la práctica. También indican que estas variables deben considerarse cuidadosamente a la hora de implementar procesos formativos que contribuyan a mejorar los sistemas de evaluación del profesorado.

De la revisión de la literatura realizada se deriva que los profesores en formación y en activo tienen ideas diferentes sobre la forma de evaluar, que están relacionadas con su propia visión acerca de la enseñanza y el aprendizaje. En los profesores en activo

influye mucho su experiencia docente y el contexto educativo donde actúan, mientras que en los profesores en formación las concepciones previas parecen estar más influenciadas por el recuerdo de cómo han sido evaluados cuando eran alumnos (Hernández, 2010). Aunque hay muchos antecedentes sobre esta temática observamos que la mayoría de los estudios revisados se han realizado en el extranjero o en un contexto diferente al nuestro, siendo todavía escasos los estudios sobre este tema desarrollados en el marco que ofrece el máster de profesorado de secundaria. También echamos de menos investigaciones que profundicen en la relación entre concepciones previas sobre los diversos procesos educativos y los modelos didácticos subyacentes a tales concepciones (Solís et al., 2012) o estudios que aporten datos sobre las fortalezas y debilidades del proceso de formación desarrollado en cada experiencia (Perales, 2014). En este nuevo contexto nos parece necesario desarrollar estudios sobre la evaluación del aprendizaje que traten de llevar a la práctica el enfoque reflexivo en la formación inicial docente (García Carmona 2013), describiendo el tipo de actividades que se realizan en el aula y analizando las concepciones de los estudiantes del máster que salen a la luz en el desarrollo de tales actividades. La necesidad de tener en cuenta tales aspectos constituye la justificación y el principal interés del presente estudio.

3. Diseño de la investigación

En esta investigación se ha adoptado un enfoque metodológico mixto, basado en un análisis cualitativo y cuantitativo de los datos recogidos. En primer lugar intentamos acercarnos al conocimiento de las ideas del alumnado mediante técnicas de análisis cualitativo (Rodríguez, Gil y García, 1996) y en segundo lugar se realiza un análisis cuantitativo de frecuencias y porcentajes correspondientes a las diversas categorías de ideas recogidas en la investigación. Se trata, por tanto, de un estudio descriptivo, basado en un diseño no experimental.

3.1. Objetivos

El estudio se enmarca en un proyecto de investigación más amplio cuyo contexto, finalidad global y resultados previos se han expuesto en varios trabajos anteriores relacionados con las concepciones de los futuros docentes sobre los procesos de enseñanza y aprendizaje (Pontes et al., 2013, 2015). En esta ocasión nos centraremos en el análisis del pensamiento docente de los estudiantes del MPES sobre la evaluación, persiguiendo los objetivos específicos siguientes:

- ✓ Explorar las concepciones previas de los estudiantes del MPES sobre la evaluación del aprendizaje de la ciencia.
- ✓ Analizar sus ideas sobre los recursos e instrumentos de evaluación que consideran más importantes para valorar el proceso de aprendizaje del alumnado de secundaria.
- ✓ Estudiar el grado de coherencia interna o consistencia inter-contextual de los modelos didácticos subyacentes en las concepciones previas de los participantes sobre diversos aspectos de la evaluación.

3.2. Participantes

En esta experiencia formativa han participado 71 estudiantes del MPES de diversas especialidades del área científico-técnica, que han cursado el citado máster en la Universidad de Córdoba (España), entre los cursos académicos 2009-10 y 2010-11, de

los que 39 eran mujeres y 32 eran hombres, con una edad media de 26.5 años. Por especialidades han participado 2 grupos de Biología, Salud y Deporte (28 estudiantes), 2 grupos de Física-Química (19 estudiantes), 1 grupo de Tecnología (13 estudiantes) y 1 grupo de Matemáticas e Informática (11 estudiantes). Algo menos de una quinta parte de tales estudiantes han tenido experiencia docente y todos han expresado su deseo de cursar el máster de profesorado por varias razones analizadas en un estudio previo: algo más de la mitad por tener acceso laboral a centros docentes públicos y privados, una cuarta parte han mostrado un interés de tipo vocacional por la docencia y el resto alude a razones de otro tipo, como ampliar el currículum académico o ampliar el campo de posibles salidas profesionales. En cualquier caso todos los participantes están bastante interesados en adquirir una buena formación pedagógica y didáctica, sobre todo en aspectos metodológicos, comunicativos y en el uso de recursos docentes (Pontes y Poyato, 2014).

3.3. Recogida y análisis de datos

La recogida de datos de este estudio se realizó en una clase del módulo específico del máster, utilizando una batería de cuestiones abiertas que formaban parte del programa de actividades de aula y que se exponen en el cuadro 1. Para ello hemos contado con la colaboración de diversos docentes de la asignatura “*Aprendizaje y Enseñanza de las Materias de la Especialidad (AEME)*”, que han utilizado el material didáctico elaborado para este proyecto. Estas cuestiones se han planteado antes de iniciar el bloque de contenidos sobre evaluación incluido en dicha materia, de modo que las concepciones registradas en este estudio pueden considerarse como ideas previas de los estudiantes del MPES, aunque tales ideas pueden estar influenciadas por los conocimientos adquiridos en las materias de formación psicopedagógica cursadas en el módulo genérico del máster.

A lo largo del proyecto se ha aplicado la misma metodología de análisis cualitativo empleada en trabajos anteriores (Pontes et al., 2015), basada en la separación y la clasificación de ideas primarias en diferentes categorías y subcategorías a partir de las respuestas de los estudiantes en cada cuestión (Rodríguez et al., 1996; Buendía et al., 1999; Hernández, 2010). En esta fase de análisis han intervenido tres investigadores expertos en didáctica de las ciencias y formación del profesorado, adoptando criterios comunes a la hora de categorizar y clasificar las ideas básicas que forman parte de cada respuesta. Este método de trabajo es dificultoso y no está exento de contradicciones, pero por coherencia metodológica hemos preferido mantener el mismo enfoque en el tratamiento de los datos recogidos en esta primera fase del proyecto. En una etapa posterior hemos comenzado a utilizar el software de análisis cualitativo “Atlas.ti” al explorar otros aspectos del pensamiento docente relacionados con el desarrollo de la identidad profesional (Serrano y Pontes, 2016). Tras el proceso de categorización anterior se han contabilizado las ideas básicas que forman parte de cada respuesta y se ha realizado un análisis de frecuencias y porcentajes que se incluyen en cada una de las tablas de ideas registradas en las diferentes cuestiones. Por limitaciones de espacio, en este estudio sólo se analizarán los resultados correspondientes a las actividades A3 y A4, que son las que han proporcionado un mayor número de ideas interesantes sobre la evaluación, por parte de los participantes.

En tales cuestiones se han analizado sus opiniones y se han registrado las ideas expresadas individualmente por escrito, encontrando en todos los casos un número mayor de ideas básicas (y relativamente diferenciadas dentro de una respuesta más o

menos amplia) que el número de sujetos encuestados, porque cada estudiante ha expresado por término medio más de una idea diferente en cada cuestión. Por ello el análisis de frecuencias relativas o porcentajes que se mostrará en la sección de resultados se realizará sobre el cómputo total de ideas diferentes registradas en cada cuestión. Sin embargo, a la hora de realizar un estudio de consistencia de los modelos didácticos subyacentes en tales concepciones se utilizará sólo la idea principal expresada por cada sujeto en ambas cuestiones, de modo que en dicho análisis los porcentajes se refieren al número de ideas principales, que a su vez coinciden con el número de participantes en la experiencia.

- A1. Cita por orden de importancia las principales competencias (máximo 5 o 6) que debe adquirir el profesorado de enseñanza secundaria de tu especialidad durante el proceso de formación inicial, incluyendo si procede la competencia evaluadora (saber evaluar).*
- A2. ¿Cuáles son los principales fines o funciones de la evaluación educativa en el área de ciencias y en el contexto de la educación secundaria?*
- A3. ¿En qué consiste la evaluación educativa y qué relación existe entre evaluación del aprendizaje y mejora de la enseñanza?*
- A4. ¿Qué recursos o qué instrumentos de evaluación crees que son los más importantes para valorar el proceso de aprendizaje del alumnado en materias de tu especialidad?*
- A5. ¿Cómo y cuando es necesario evaluar? Indicar posibles etapas del proceso de evaluación comentando sus características.*
- A6. ¿Quiénes son los agentes del proceso de evaluación? ¿Es necesario evaluar también el proceso de enseñanza? ¿Por qué motivos y cómo hacerlo?*

Figura 1: Cuestiones abiertas sobre la evaluación

Fuente: Elaboración propia.

4. Resultados

En primer lugar se analizan las principales concepciones previas de los participantes sobre las cuestiones planteadas en las actividades A3 y A4 del programa-guía recogido en el Cuadro 1, mostrando las tablas de resultados que se han obtenido en tales cuestiones, tras el análisis de las respuestas correspondientes a cada cuestión y la categorización de las ideas registradas en cada pregunta. Después de cada tabla se incluyen y comentan ejemplos de algunos tipos de respuestas representativas en cada categoría de ideas. Al tratarse de cuestiones generales hemos observado ciertas pautas o tendencias que son comunes a los estudiantes de diferentes especialidades (Biología, Física, Matemáticas o Tecnología), y son tales pautas las que se reflejan principalmente en los datos cuantitativos que se muestran en las tablas posteriores. Sin embargo, al tratar de analizar las ideas sobre evaluación en conexión con los modelos didácticos subyacentes del alumnado del MPES, hemos sido conscientes de que puede haber conexión entre tales modelos y las concepciones epistemológicas de los estudiantes, que a su vez dependen de la formación académica previa y del conocimiento didáctico del contenido de la propia disciplina (Valbuena, 2007). Por ello creemos que sería conveniente realizar estudios posteriores más específicos, que permitan profundizar en este asunto, ya que en este estudio sólo se han abordado aspectos muy generales del pensamiento docente sobre la evaluación.

4.1. La evaluación educativa y su relación con la mejora de la enseñanza

En la cuestión planteada en la actividad A3 los futuros docentes deben explicar en qué consiste la evaluación educativa y qué relación existe entre evaluación y mejora de la educación científico-técnica. Se han analizado las opiniones de los participantes sobre esta temática y se han registrado 134 ideas sobre este tema, lo cual supone una media de 1,9 por sujeto, identificando diferentes categorías y sub-categorías de opiniones, que han permitido realizar un estudio de frecuencias y porcentajes, cuyos datos se muestran en la tabla 1 y se comentan posteriormente.

En la categoría A se recogen opiniones sobre la evaluación que podemos considerar próximas al enfoque educativo tradicional según la clasificación propuesta por diversos autores (Porlán et al., 1997), ya que se identifica principalmente con aspectos como la medida de los conocimientos adquiridos por los alumnos, lo cual se refleja en una nota de examen o calificación numérica y que, en la práctica, actúa como un mecanismo de selección o de promoción. En esta categoría se incluye algo más de una tercera parte de las ideas registradas. Por ejemplo:

La evaluación supone medir el nivel de conocimientos alcanzado por los alumnos respecto a los objetivos previstos". "En la enseñanza secundaria no es posible llevar un seguimiento individual del aprendizaje de cada alumno, de modo que hay que hacer exámenes escritos, porque de lo contrario los alumnos no estudiarían la asignatura.

En la categoría B se incluyen opiniones que recogen una visión más integral de la evaluación y que podemos considerar próximas al modelo educativo socio-constructivista (Fuentes et al., 2009). Así, se asume que la evaluación supone un proceso en el que se pueden recoger múltiples datos sobre el proceso de aprendizaje y que tales datos sirven de instrumento para la reflexión del profesorado y para favorecer la mejora de la acción docente. En esta categoría se incluye algo menos de una tercera parte de las respuestas analizadas. Por ejemplo:

La evaluación es el proceso mediante el cual se analiza el grado de consecución de las metas educativas que te has propuesto en la asignatura. La relación con la mejora de la enseñanza es evidente porque la evaluación siempre se realiza con la intención de mejorar el proceso de enseñanza y aprendizaje, de conocer dónde hay deficiencias para resolverlas o si es necesario cambiar la metodología, etc.

En la categoría C se recogen otras opiniones sobre la evaluación que podríamos considerar en una zona intermedia entre el enfoque tradicional y el enfoque socio-constructivista, en la que se integran algo más de una cuarta parte de las ideas registradas:

Evaluar supone una valoración del alumnado en la adquisición de los conocimientos mínimos. Pero la mayoría de veces nos centramos en qué se enseña y no en cómo se enseña, ahí es donde está la clave para mejorar. Siempre es necesario evaluar el nivel que adquieren los alumnos mediante exámenes. Pero también es importante indagar lo que de verdad aprenden los alumnos y reflexionar sobre aquello que favorece la mejora de la educación.

Como ocurría en otras cuestiones analizadas en este proyecto relacionadas con el aprendizaje en materias del área científico-técnica (Pontes et al., 2015), también en torno a la evaluación aparecen numerosas concepciones próximas al enfoque educativo tradicional (A), pero es interesante que haya casi un tercio de opiniones de carácter innovador (B), próximas al enfoque constructivista (Solís et al., 2012). Por otra parte observamos que en la posición intermedia o ambivalente (B) subyace la idea de que la

evaluación es un proceso de valoración del rendimiento escolar del alumno, pero también se incorporan otras ideas próximas al enfoque innovador como son la valoración de competencias y destrezas, la evaluación de conocimientos previos para valorar el avance del alumno o la posibilidad de incorporar a los alumnos al proceso de evaluación, como complemento a la valoración del profesor.

Tabla 1. Concepciones sobre la evaluación educativa

TIPOS DE IDEAS REGISTRADAS	FREC.	%
(A) Enfoque tradicional:	46	(34,3)
▪ La evaluación como medida de conocimientos adquiridos	24	17,9
▪ La evaluación como sinónimo de nota o calificación	16	11,9
▪ La evaluación como mecanismo de selección	6	4,5
(B) Enfoque innovador:	42	(31,4)
▪ La evaluación como recogida de múltiples datos sobre el proceso de aprendizaje (conocimientos, actitudes, destrezas, esfuerzo,...)	21	15,7
▪ La evaluación como instrumento de reflexión del profesorado	12	9,0
▪ La evaluación como recurso de indagación y mejora de la acción docente	9	6,7
(C) Enfoque intermedio	37	(27,6)
▪ Medir conocimientos adquiridos y otros aspectos (competencias, destrezas,...)	19	14,2
▪ Medir para comprobar el rendimiento del proceso educativo, analizando lo que saben los alumnos antes y después.	8	5,9
▪ Coherencia entre método de enseñanza y sistema de evaluación	6	4,5
▪ Incorporar a los alumnos al proceso de evaluación, como complemento a la valoración del profesor	4	3,0
(D) Otras ideas poco explícitas (difíciles de categorizar)	8	(5,9)

Fuente: Elaboración propia.

4.2. Recursos e instrumentos para evaluar el aprendizaje en materias de la especialidad

En la cuestión planteada en la actividad A4 se solicita a los estudiantes del máster FPES que expongan sus ideas sobre los recursos o los instrumentos de evaluación que consideran más importantes para valorar el proceso de aprendizaje del alumnado de secundaria en materias del área científico-técnica. Tras analizar las 124 respuestas distintas encontradas para esta cuestión (1,7 respuestas por alumno), se han identificado diferentes categorías y se ha realizado un estudio de frecuencias y porcentajes, cuyos datos se muestran en la tabla 2 y se comentan a continuación.

En los resultados correspondientes a esta cuestión también se recogen en la categoría A las opiniones sobre los recursos e instrumentos de evaluación relacionadas con el enfoque educativo tradicional (Oliva, 2008), ya que el principal medio utilizado para medir el rendimiento académico es el examen (aunque incluya diversos tipos de preguntas tales como cuestiones, ejercicios, etc.) y en otros casos se mantiene el examen como recurso básico para obtener una calificación, pero se complementa (de forma vaga e imprecisa) con otro tipo de datos: la participación en clase, las tareas realizadas en casa, etc. En esta categoría se incluye alrededor de una sexta parte de las ideas registradas sobre este tema. Por ejemplo:

La evaluación está orientada a medir los conocimientos adquiridos por cada alumno y, por tanto, creo que el examen escrito (incluyendo teoría y problemas) debe ser el principal recurso. Veo difícil que en asignaturas como la biología se puedan sustituir los exámenes

por otro tipo de instrumentos de evaluación, pero se puede complementar la nota del examen con otros datos.

En la categoría B se incluyen ideas sobre la forma de llevar a cabo la evaluación que podríamos considerar próximas al modelo educativo constructivista, porque en muchos casos se asume la evaluación como un proceso continuo e integral, que requiere el uso de una amplia gama de recursos e instrumentos destinados a evaluar diferentes aspectos del proceso de aprendizaje (conocimientos teóricos, destrezas de tipo práctico, capacidad de análisis y resolución de problemas, capacidad de relacionar conceptos, actitud en clase, participación y capacidad de trabajar en equipo). Se concede importancia a los conocimientos previos y a la observación del trabajo realizado por el alumno en el aula, o se habla de registrar tales observaciones en el diario del profesor. Pese a las dificultades de este modelo de evaluación, encontramos que más dos quintas partes de las ideas registradas pueden incluirse en esta categoría. Por ejemplo:

El aprendizaje de las ciencias es un proceso complejo y se deben usar instrumentos de evaluación variados y complementarios como explicación de hechos, resolución de problemas, mapas conceptuales...En Biología es importante saber evaluar las competencias generales y específicas de la materia, las destrezas mostradas en el aula taller al desarrollar proyectos, la actitud del alumno en el aula y también hay que evaluar los conocimientos adquiridos mediante diferentes tipos de pruebas escritas.

Tabla 2. Ideas sobre recursos e instrumentos de evaluación

TIPOS DE IDEAS REGISTRADAS	FREC.	(%)
(A) Evaluación orientada a medir conocimientos	20	(16,1)
- El examen escrito como principal recurso	11	(8,9)
- Complementar la nota de exámenes con otros datos	9	(7,2)
(B) La evaluación como un proceso continuo e integral	53	(42,7)
- Uso de instrumentos de evaluación variados	26	20,1
- Evaluación continua del trabajo del aula y del aprendizaje desarrollado	15	(12,1)
- Evaluación de competencias, destrezas, actitudes y conocimientos	12	(9,8)
(C) Posición ambivalente	42	(33,9)
- Importancia de evaluar aspectos prácticos	27	(21,8)
- Utilizar las nuevas tecnologías	10	(8,1)
- Incluir elementos de autoevaluación o coevaluación	5	(4,0)
(D) Otras ideas poco explícitas (difíciles de categorizar)	9	(7,3)

Fuente: Elaboración propia.

De igual forma que en la cuestión anterior, también se recogen en la categoría C otras opiniones, sobre los instrumentos de evaluación, que podrían considerarse como visiones ambivalentes o intermedias entre el modelo tradicional y el enfoque innovador. Tales opiniones conceden importancia al examen escrito como instrumento de medida del rendimiento del alumno, pero consideran necesario complementar este aspecto usando otros recursos como son las nuevas tecnologías, las prácticas de laboratorio, el desarrollo de proyectos tecnológicos o la realización de actividades de autoevaluación y coevaluación. En esta categoría hemos registrado algo más de una tercera parte de las ideas mostradas por los futuros profesores del área de ciencias:

El examen clásico es un recurso importante para evaluar el aprendizaje de cada alumno, pero se deben tener en cuenta otros datos recogidos mediante el diario del profesor o la observación del trabajo experimental de los alumnos para evaluar la parte práctica de la asignatura". "En ciencias como la física y química hay que hacer muchos ejercicios y prácticas, fomentando la participación porque así resulta más fácil aprender, con lo cual se evaluará no solo algo escrito sino también el progreso realizado día a día.

En torno a esta cuestión también aparecen diversas opiniones difíciles de categorizar (7,3 %), porque algunos de los futuros profesores emiten respuestas muy vagas ("Me

bastaría con inculcar una curiosidad científica”) sobre la forma en la que se podría llevar a cabo el proceso de evaluación en sus respectivas materias.

4.3. Consistencia de los modelos didácticos subyacentes en las concepciones previas de los futuros docentes de ciencias sobre la evaluación

En la cuestión planteada en la actividad A3 (tabla 1) los futuros docentes deben explicar en qué consiste la evaluación educativa y qué relación existe entre evaluación y mejora de la educación científico-técnica. Por otra parte, en la cuestión planteada en la actividad A4 (tabla 2) se solicita a los estudiantes del máster FPES que expongan sus ideas sobre los recursos o los instrumentos de evaluación que consideran más importantes para valorar el proceso de aprendizaje del alumnado de secundaria en las materias del área.

Sobre ambas cuestiones se han recogido diversas ideas que se han clasificado con arreglo a dos modelos de pensamiento docente bien diferenciados como son el enfoque tradicional (A) y el enfoque innovador o constructivista (B). El primero se caracteriza por una visión de los procesos educativos que pone su acento en el protagonismo del profesor y los contenidos de la enseñanza, mientras que en el segundo cobra gran importancia el protagonismo del alumno en el aprendizaje, el papel de las actividades el proceso de enseñanza y la necesidad de llevar a cabo una evaluación acorde con la visión del aprendizaje y la enseñanza. Así mismo, se han encontrado en ambas cuestiones un número importante de ideas algo menos definidas, que presentan un carácter dual o ambivalente entre los enfoques extremos (A y B) y que se han incluido en el llamado enfoque intermedio (C). En ambos casos también hay ideas muy confusas, indefinidas o difíciles de clasificar en alguno de los modelos descritos anteriormente y que hemos incluido en el grupo D, que es bastante minoritario.

Para ver hasta qué punto existe coherencia interna en los puntos de vista de los participantes sobre tales cuestiones, o en qué medida son consistentes los modelos didácticos subyacentes a las concepciones de los futuros docentes sobre la evaluación, se ha realizado un estudio cruzado de los resultados de ambas cuestiones utilizando un análisis basado en tablas de contingencia, incluyendo algunos estadísticos que miden el grado de relación entre los modelos cruzados y cuyos resultados se muestran en la tabla 3. Conviene recordar un aspecto señalado anteriormente y es que en ambas cuestiones se ha registrado, por término medio, más de una idea diferente y, por tanto, los porcentajes expresados en las tablas 1 y 2 se referían al número completo de ideas de cada cuestión, que es mayor que el número de sujetos de la muestra. Sin embargo, para estudiar la consistencia de los modelos didácticos subyacentes en las concepciones sobre evaluación, correspondientes a las dos cuestiones analizadas en este trabajo, se ha utilizado sólo la idea principal expresada por cada sujeto en ambas cuestiones. Por tanto, las frecuencias y porcentajes que se muestran en la tabla 3 se refieren tanto al número de ideas principales como al número de estudiantes, que en este caso son coincidentes.

En los resultados mostrados en la tabla 3 se puede observar que hay cierta relación de correspondencia entre los modelos de pensamiento de carácter tradicional (A), innovador (B) e intermedio (C) subyacentes a las principales concepciones de los futuros docentes sobre la evaluación del aprendizaje de las ciencias, registradas en las actividades A3 y A4. Lo cual indica que los estudiantes del MPES se identifican en cierta medida con alguno de tales enfoques didácticos y tienden a expresar ideas relativamente consistentes con su modelo de pensamiento en cuestiones diferentes. Sin

embargo, los parámetros estadísticos que miden el grado de relación de tales modelos en las dos cuestiones analizadas no son suficientemente significativos para considerar que este resultado es del todo afirmativo. Creemos que ello puede ser debido a las limitaciones propias de los análisis de tipo cualitativo y, por tanto, consideramos necesario realizar nuevos estudios sobre esta temática en los que se utilicen instrumentos y técnicas de investigación basados en un enfoque cuantitativo.

Tabla 3. Tabla de contingencia entre modelos didácticos registrados en las actividades A3 y A4

MODELOS DOCENTES EN A3		MODELOS DOCENTES EN A4				TOTAL
		A	B	C	D	
A	Frec.	6	13	5	0	24
	(%)	50,0	41,9	21,7	0,0	33,8
B	Frec.	4	11	5	2	22
	(%)	33,3	35,5	21,7	40,0	31,0
C	Frec.	1	7	11	2	21
	(%)	8,3	22,6	47,8	40,0	29,6
D	Frec.	1	0	2	1	4
	(%)	8,3	0,0	8,7	20,0	5,6
Total	Frec.	12	31	23	5	71
	(%)	16,9	43,7	32,4	7,0	100,0

Notas: Razón de verosimilitudes: 17,453 (p: 0,046)

Tau-b de Kendall: 0,322 (p: 0,094)

Fuente: Elaboración propia.

5. Discusión y conclusiones

En este trabajo hemos mostrado los resultados de una experiencia educativa realizada durante el proceso de formación inicial del profesorado de secundaria, en diversas especialidades del área científico-técnica, tratando de poner en práctica el enfoque reflexivo en la formación docente (Hewson, 1993; García Carmona, 2013) y realizando en el aula actividades que permiten explicitar las ideas previas sobre la evaluación del aprendizaje (Buendía et al., 1999; Segers y Tillema, 2011). En primer lugar se han analizado las concepciones de los participantes sobre la evaluación del aprendizaje y sus ideas sobre los recursos e instrumentos de evaluación que consideran más importantes para valorar el proceso de aprendizaje del alumnado de secundaria. También hemos tratado de relacionar tales ideas con los modelos didácticos que presentan mayor consenso en la literatura sobre el pensamiento del profesorado (Oliva, 2008; Hernández y Maquilón, 2010; Solís et al., 2012), decantándonos por una clasificación de ideas que se mueve en torno a tres categorías principales: un modelo tradicional de enseñanza y evaluación, bien diferenciado de un modelo didáctico que hemos denominado innovador y que se encuentra próximo a la perspectiva socio-constructivista, existiendo entre ambos polos un enfoque intermedio de carácter dual o ambivalente, hecho que también pudimos constatar al analizar las concepciones de los futuros docentes sobre los procesos de aprendizaje (Pontes et al., 2015).

En la primera cuestión analizada hemos encontrado una proporción similar de concepciones previas sobre la evaluación que corresponden, respectivamente, al enfoque educativo tradicional, al enfoque innovador y al enfoque intermedio o dual. Sin embargo, en la segunda cuestión, al analizar las ideas sobre el uso de recursos e instrumentos, encontramos una proporción mayor de ideas innovadoras que ideas

tradicionales. Nos parece interesante que muchos de los participantes muestren opiniones sobre la manera concreta de evaluar el aprendizaje de la ciencia que resultan próximas a posiciones constructivistas o alternativas (Porlán et al., 1997), aunque también hay un porcentaje importante de ideas de tipo tradicional. Nos parece necesario, por tanto, seguir avanzando y profundizando en el diseño de procesos formativos orientados a fomentar una visión de la evaluación de carácter integral que tenga en cuenta, además de los conocimientos adquiridos por los alumnos, la actitud hacia el aprendizaje, las destrezas o competencias adquiridas, el esfuerzo realizado y el comportamiento en aula. De esta forma se podría conseguir que las ideas sobre la evaluación de carácter tradicional se transformen en concepciones coherentes con un modelo didáctico constructivista e innovador (Solís et al., 2012).

En segundo lugar, pese a tratarse esencialmente de una investigación de carácter exploratorio y cualitativo, hemos tratado de estudiar también el grado de coherencia interna o consistencia inter-contextual de los modelos didácticos subyacentes en las concepciones previas de los participantes sobre los aspectos analizados anteriormente. El hecho de encontrar porcentajes diferentes entre las ideas de corte tradicional e innovador sobre la evaluación en función del tipo de cuestión planteada, según se aprecia en las tablas 1 y 2, indica que las concepciones sobre la evaluación de los futuros docentes no son del todo consistentes o que dependen del contexto del problema (Wang et al., 2010). Por ello hemos introducido algunas técnicas propias del análisis cuantitativo para realizar un estudio cruzado de los resultados de las dos primeras cuestiones, utilizando un análisis basado en tablas de contingencia, cuyos resultados no pueden considerarse concluyentes por las limitaciones metodológicas que se han comentado al exponer los resultados mostrados en la tabla 3. En efecto, hemos observado que hay cierta relación de correspondencia entre los modelos de pensamiento de carácter tradicional, innovador e intermedio sobre la evaluación del aprendizaje, pero los parámetros estadísticos que miden el grado de relación de tales modelos no son suficientemente significativos, lo cual puede deberse al hecho de que tales resultados proceden de cuestiones abiertas relativamente diferentes y a las limitaciones propias del análisis cualitativo (Dixon y Haigh, 2009). También conviene considerar que las ideas sobre evaluación analizadas en este estudio corresponden al nivel declarativo cuya categorización resulta en ocasiones difícil (Valbuena, 2007; Pontes et al., 2013). Por tanto, creemos que para profundizar en el estudio de la consistencia interna de las concepciones sobre la evaluación de los futuros docentes sería necesario realizar nuevas investigaciones en las que se utilicen instrumentos y técnicas de carácter cuantitativo.

Como conclusión principal, podemos destacar que las concepciones de los estudiantes del MPES sobre la evaluación se pueden relacionar principalmente con los dos enfoques educativos antagónicos, de carácter tradicional o de carácter innovador (próximo al modelo constructivista), y con otro enfoque intermedio que se deriva de aquellas opiniones en las que aparecen elementos de ambos (Turpo-Gebera, 2011; Halinen et al., 2014). Un número ligeramente mayor de opiniones se relaciona con el enfoque innovador o centrado en el alumno, lo cual nos parece interesante como punto de partida para la formación posterior. Pero también encontramos contradicciones y diferencias de enfoque según la cuestión abordada en cada momento y, por otra parte, pensamos que en esta fase formativa los futuros docentes utilizan una especie de constructivismo idealizado para responder a diversas cuestiones, sin ser plenamente conscientes de la dificultad que supone trasladar este enfoque a la práctica educativa real (Mellado et al., 1999; Contreras, 2010; Pontes et al., 2015). Creemos, en cualquier

caso, que los resultados de este estudio pueden servir como base para avanzar posteriormente en el estudio del pensamiento docente sobre la evaluación utilizando nuevos instrumentos, que permitan recoger datos de carácter cuantitativo y aplicar técnicas estadísticas de análisis multivariante, con vistas a estudiar la consistencia interna de los modelos didácticos y su evolución a través del proceso de formación inicial docente. En la segunda etapa de este proyecto de investigación trataremos de abordar estos aspectos y también trataremos de recoger datos sobre las opiniones del alumnado del MPES acerca del proceso formativo desarrollado en el aula.

Referencias

- Abell, S. K. (2007). Research on science teacher knowledge. En S. K. Abell y G. Lederman (Eds.), *Handbook of research on science education* (pp. 1105-1149). Londres: Lawrence Erlbaum Associates.
- Borko, H. (1997). New forms of classroom assessment: Implications for staff development. *Theory into Practice*, 36(4), 231-238. doi:10.1080/00405849709543773.
- Brown, G. T., Lake, R. y Matters, G. (2011). Queensland teachers' conceptions of assessment: The impact of policy priorities on teacher attitudes. *Teaching and Teacher Education: An International Journal of Research and Studies*, 27(1), 210-220. doi:10.1016/j.tate.2010.08.003.
- Buendía, L., Carmona, M., González, D. y López, R. (1999). Concepciones de los profesores de educación secundaria sobre evaluación. *Educación XX1*, 2(1), 125-154. doi:10.5944/educxx1.2.0.377.
- Carvajal, E. y Gómez, M. (2002). Concepciones y representaciones de los maestros de secundaria y bachillerato sobre la naturaleza, el aprendizaje y la enseñanza de las ciencias. *Revista Mexicana de Investigación Educativa*, 7(16), 577-602.
- Contreras, A. (2010). *Las creencias y actuaciones curriculares de los profesores de ciencias de Secundaria de Chile*. (Tesis doctoral). Universidad Complutense de Madrid.
- De Vries, W. (2015). Las múltiples facetas de la evaluación educativa. *Revista Mexicana de Investigación Educativa*, 20(66), 679-683.
- Dixon, H. y Haigh, M. (2009). Changing mathematics teachers' conceptions of assessment and feedback. *Teacher Development*, 13(2), 173-186. doi:10.1080/13664530903044002.
- Fuentes, M. J., García, S. y Martínez, C. (2009). ¿En qué medida cambian las ideas de los futuros docentes de Secundaria sobre qué y cómo enseñar, después de un proceso de formación? *Revista de Educación*, 349, 269-294.
- García-Carmona, A. (2013). Educación científica y competencias docentes: Análisis de las reflexiones de futuros profesores de Física y Química. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 10(Núm. Extraordinario), 552-567.
- Halinen, K., Ruohoniemi, M., Katajavuori, N. y Virtanen, V. (2014). Life science teachers' discourse on assessment: A valuable insight into the variable conceptions of assessment in higher education. *Journal of Biological Education*, 48(1), 16-22. doi:10.1080/00219266.2013.799082.
- Hernández, F. y Maquilón, J. (2010). Las concepciones de la enseñanza. Aportaciones para la formación del profesorado. *Revista Electrónica Interuniversitaria de Formación del Profesorado*, 13(3), 17-25.
- Hernández, L. M. (2010). Evaluar para aprender: Hacia una dimensión comunicativa, formativa y motivadora de la evaluación. *Enseñanza de las ciencias*, 28(2), 285-290.

- Hewson, P. W. (1993). Constructivism and reflective practice in science teacher education. En M. L. Montero y J. M. Vez (Eds.), *Las didácticas específicas en la formación del profesorado* (pp. 259-275). Santiago de Compostela: Tórculo.
- Meirink, J., Meijer, P., Verloop, N. y Bergen, T. (2009). Understanding teacher learning in secondary education: The relations of teacher activities to changed beliefs about teaching and learning. *Teaching and teacher education*, 25(1), 89-100. doi:10.1016/j.tate.2008.07.003.
- Mellado, V., Blanco, L. y Ruiz, C. (1999). *Aprender a enseñar ciencias experimentales en la formación inicial de profesorado*. Badajoz: ICE de la Universidad de Extremadura.
- Monteagudo Fernández, J., Molina Puche, S. y Miralles Martínez, P. (2015). Opiniones sobre evaluación de los profesores de segundo ciclo de ESO de Geografía e Historia en España: El caso de la Región de Murcia. *Revista Mexicana de Investigación Educativa*, 20(66), 737-761.
- Oliva, J. M. (2008). *Metodología y recursos educativos: Diseño de materiales didácticos y actividades de aprendizaje*. En A. Pontes [Coord.], *Aspectos generales de la formación psicopedagógica del profesorado de educación secundaria*. (pp. 193-216). Córdoba: Servicio de Publicaciones de la UCO.
- Perales, F. J. (2014). La reforma de la formación inicial del profesorado de ciencias de secundaria: Propuesta de un diseño del currículo basado en competencias. *Enseñanza de las Ciencias*, 32(1), 9-28. doi:10.5565/rev/ensciencias.898.
- Pontes, A. y Poyato, F. J. (2014). El enfoque reflexivo en la formación inicial del profesorado de secundaria: Motivaciones por la docencia. *Cuadernos de Pedagogía*, 445, 34-37.
- Pontes, A., Poyato, F. J. y Oliva, J. M. (2015). Concepciones sobre el aprendizaje en estudiantes del máster de profesorado de enseñanza secundaria del área de ciencia y tecnología. *Profesorado: Revista de Currículum y Formación del Profesorado*, 19(2), 225-243.
- Pontes, A., Serrano, R. y Poyato, F. J. (2013). Concepciones y motivaciones sobre el desarrollo profesional docente en la formación inicial del profesorado de enseñanza secundaria. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 10(Núm. Extraordinario), 533-551. doi:10498/15612.
- Pontes, A., Serrano, R., Muñoz, J. M. y López, I. (2011). Innovación educativa sobre aprendizaje colaborativo con Cmap Tools en la formación inicial docente. *Revista Iberoamericana de Evaluación Educativa*, 4(2), 136-154.
- Porlán, R. y Martín, R. (2006). ¿Cómo progresa el profesorado al investigar problemas prácticos relacionados con la enseñanza de la ciencia? *Alambique: Didáctica de las Ciencias Experimentales*, 48, 92-99.
- Porlán, R., Rivero, A. y Martín del Pozo, R. (1997). Conocimiento profesional y epistemología de los profesores (I): Teoría, métodos e instrumentos. *Enseñanza de las Ciencias*, 15(2), 155-171.
- Prieto, M. y Contreras, G. (2008). Las concepciones que orientan las prácticas evaluativas de los profesores: Un problema a develar. *Estudios Pedagógicos*, 34(2), 245-262. doi:10.4067/s0718-07052008000200015.
- Remesal, A. (2011). Primary and secondary teachers' conceptions of assessment: A qualitative study. *Teaching and Teacher Education: An International Journal of Research and Studies*, 27(2), 472-482. doi:10.1016/j.tate.2010.09.017.
- Rodríguez, G., Gil, J. y García, E. (1996). *Metodología de la investigación cualitativa*. Málaga: Aljibe.

- Segers, M. y Tillema, H. (2011). How do Dutch secondary teachers and students conceive the purpose of assessment? *Studies in Educational Evaluation*, 37(1), 49-54. doi:10.1016/j.stueduc.2011.03.008.
- Serrano, R. y Pontes, A. (2015). Nivel de desarrollo de las competencias y objetivos generales del Máster formación del profesorado de enseñanza secundaria. *Perfiles Educativos*, 37(150), 39-55.
- Serrano, R. y Pontes, A. (2016). Representación del conocimiento sobre la identidad profesional docente en la formación inicial del profesorado. En R. Serrano, M. E. Gómez y C. A. Huertas (Eds), *La educación sí que importa en el siglo XXI*. (pp. 115-128). Madrid: Síntesis.
- Sethusha, M. J. (2013). A vision of improvement of learning: South African teachers' conceptions of classroom assessment. *Perspectives in Education*, 31(2), 14-21.
- Solís, E., Porlán, R. y Rivero, A. (2012). ¿Cómo representar el conocimiento curricular de los profesores de ciencias y su evolución? *Enseñanza de las Ciencias*, 30(3), 9-30.
- Turpo-Gebera, O. W. (2011). Concepciones y prácticas evaluativas de los docentes del área curricular de ciencias en las instituciones de enseñanza públicas de educación secundaria. *Revista Iberoamericana de Evaluación Educativa*, 4(2), 213-233.
- Valbuena, E. O. (2007). *El conocimiento didáctico del contenido biológico. Estudio de las concepciones disciplinares y didácticas de futuros docentes de la Universidad Pedagógica Nacional (Colombia)*. (Tesis doctoral). Universidad Complutense de Madrid.
- Vergara, C. E. (2012). Análisis de las concepciones de evaluación del aprendizaje de docentes destacados de Educación Básica. *Revista Iberoamericana de Evaluación Educativa*, 5(3), 249-273.
- Wang, J. R., Kao, H. L. y Lin, S. W. (2010). Preservice teachers' initial conceptions about assessment of science learning: The coherence with their views of learning science. *Teaching and Teacher Education: An International Journal of Research and Studies*, 26(3), 522-529. doi:10.1016/j.tate.2009.06.014.

El Papel de la Evaluación a Gran Escala como Política de Rendición de Cuentas en el Sistema Educativo Mexicano

The Role of the Large-Scale Assessment as Political Accountability in the Mexican Educational System

José Alfonso Jiménez Moreno*

Universidad Nacional Autónoma de México

En la actualidad, la evaluación a gran escala que se realiza mediante instrumentos estandarizados de aplicación nacional e internacional es una de las herramientas más utilizadas para el conocimiento y monitoreo del estatus de los sistemas educativos. Los resultados que este tipo de evaluación arrojan suelen ser esperados por los diversos actores educativos, dado que son un insumo importante para conocer los avances de los estudiantes. Con la finalidad de comprender de manera particular el papel de este tipo de evaluaciones dentro de las políticas educativas en México, el presente artículo hace uso de datos teóricos e históricos para analizar la relación entre las políticas económicas de las últimas décadas a nivel internacional y el uso de la evaluación a gran escala como mecanismo de rendición de cuentas dentro del sistema educativo mexicano. Derivado del análisis presentado, se concluye que el uso de este tipo de evaluación por parte del sistema educativo mexicano se da como consecuencia de una constante búsqueda de competitividad económica internacional.

Palabras clave: Evaluación basada en el rendimiento, Política educativa, Rendición de cuentas, Evaluación internacional, Pruebas nacionales de competencia.

At present, the large-scale assessment that is performed by standardized instruments of national and international administration is one of the most relevant tools for understanding and monitoring the status of the educational systems. The results that this kind of assessment release are usually expected by a diversity of stakeholders, as those results are an important input to understand the progress of students. In order to understand the role of this kind of assessment in educational policy in México, this article makes use of theoretical and historical data to analyze the relationship between the economic policies of recent decades at the international level and the use of large-scale assessment as a mechanism of accountability within the Mexican education system. Derived from the presented analysis, it is concluded that the use of this type of evaluation by the Mexican educational system is given as a result of a constant search for international economic competitiveness.

Keywords: Performance based assessment, Educational policy, Accountability, International assessment, National competency tests.

1. Introducción

La evaluación a gran escala es una de las formas de evaluación educativa de mayor uso en los sistemas educativos y, técnicamente hablando, es caracterizada particularmente por su alta afinidad con la psicometría como parte de su metodología de desarrollo. Sin embargo, no es su único elemento definitorio, ya que en la práctica cotidiana dentro de los sistemas educativos sus resultados suelen ser utilizados como herramienta de decisión en materia de política educativa y, en algunos casos, como proveedora de información para toma de decisiones que involucran el ingreso o egreso de estudiantes a un nivel educativo en particular.

Además de la relevancia académica y social que este tipo de evaluación puede tener, su importancia como herramienta definitoria de políticas educativas es innegable. Frente a ello, el presente artículo tiene por objeto mostrar el papel que la evaluación a gran escala ha tenido como herramienta de consolidación de la política educativa mexicana basada en la competitividad y rendición de cuentas, para ello se realiza un análisis del papel de este tipo de evaluación en la consolidación de la política educativa influida por los procesos de globalización económica, mediante el uso de datos teóricos e históricos relacionados con los antecedentes más representativos de este tipo de evaluación, así como un recuento y una crítica de las políticas educativas que han sido determinantes en el uso de la evaluación a gran escala como herramienta de consolidación de una política neoliberal en México.

2. Antecedentes y conceptualización

La evaluación a gran escala basa su quehacer en la teoría de la medición psicológica (Martínez, 2001). Es a partir del siglo XX que su metodología comienza a consolidarse mediante el uso de elementos estadísticos como confiabilidad y error de medida (Martínez Arias, 1995, citado en Martínez, 2001). El uso de pruebas psicométricas centradas en la educación inició bajo la intención de medir el aprendizaje como rasgo psicológico, ya que se consideraba que este era el indicador que mejor podía reflejar los logros educativos, dado que mostraba de manera directa la influencia que los programas podían tener sobre los estudiantes. La evaluación a gran escala que comenzó a desarrollarse desde el siglo XX, particularmente en Estados Unidos, tenía como interés identificar si los logros de los estudiantes eran acordes con los mínimos aceptables (Martínez y Santos, 2009). De manera concreta, la evaluación a gran escala consiste en la aplicación de pruebas a una gran cantidad de alumnos, ello con la finalidad de identificar el nivel de aprendizaje logrado por un país o región (Martínez, 2008b).

La evaluación a gran escala se fundamenta en la medición del logro del aprendizaje bajo el respaldo de técnicas estadísticas. Esto ha favorecido la obtención de evidencia empírica objetiva, aspecto de gran aprecio por parte de los actores involucrados con la toma de decisiones en educación debido a la confianza que el uso de métodos objetivos trae para la definición de políticas y acciones educativas. De esta forma, el uso de técnicas psicométricas y estadísticas ha facilitado el desarrollo de la evaluación, ya que conjuntan un proceso metodológico que privilegia la objetividad con una forma eficiente de procesamiento de datos. Actualmente la metodología de la evaluación a gran escala permite incluso estimar las habilidades de una persona en función de un manejo probabilístico de los datos (Martínez, 2001).

El objetivo de la evaluación a gran escala, de acuerdo con Ravela (2006) es que permite “conocer en qué medida están siendo logrados los conocimientos y competencias que se espera que los estudiantes hayan adquirido cuando finalizan un nivel o ciclo educativo” (p. 18). Según este autor, existen diversas finalidades de una evaluación a gran escala. La primera de ellas es la certificación, que se centra en la identificación de los conocimientos y competencias estipulados para un nivel educativo en particular. Un segundo objetivo es la selección, que se utiliza con fines de ingreso a instituciones educativas. Otra de sus metas es la toma de decisiones “blandas”, que refiere al sentido formativo de la evaluación. El autor las llama “blandas” debido a que su uso no conlleva una consecuencia formal. En contraparte se encuentra el objetivo de toma de decisiones “duras”, que se relaciona con las evaluaciones que traen como consecuencia la continuidad de un proyecto o programa. Finalmente, Ravela considera el establecimiento de incentivos y la rendición de cuentas y responsabilización, que permiten valorar el avance de las acciones educativas de mejora.

La clasificación realizada por Ravela refleja a grandes rasgos las diferentes formas de comprender los objetivos de la evaluación a gran escala. Sin embargo, no representa una clasificación rígida con categorías excluyentes entre sí, dado que –de acuerdo con esta categorización– una evaluación cuyos resultados sean condicionante para que los estudiantes puedan egresar del nivel licenciatura o tengan posibilidades de obtener un título profesional puede ser considerada como de certificación, de toma de decisiones *duras* o, incluso, como elemento de rendición de cuentas para efectos de los indicadores que las instituciones utilizan dentro de sus procesos de acreditación.

Por su parte, Martínez y Santos (2009) clasifican a las evaluaciones a gran escala en dos tipos: a) Selección: Se aplican a aspirantes de ingreso a una institución educativa y centran su atención en la validez predictiva respecto al posible desempeño de los aspirantes; y b) Aprovechamiento: Centradas en los alumnos que terminan un ciclo o grado escolar y en el aseguramiento de la validez y los contenidos con una fuerte carga curricular.

Considerando los elementos que se han descrito, los aspectos definitorios de la evaluación a gran escala están supeditados por cuestiones técnicas definidas por la psicometría y los usos que se dan a los resultados. Ahora bien, cabe aclarar una precisión terminológica: en el idioma inglés a la evaluación a gran escala se la conoce como *large scale assessment*; sin embargo, en dicho idioma existe una diferencia conceptual entre *assessment* y *evaluation*, lo cual permite diferenciar al *assessment* como forma de medición u observación constante de un fenómeno en particular de la *evaluation* como elemento de retroalimentación a partir de las mediciones observadas. En castellano, se hace uso del término *evaluación* sin diferencia respecto al *assessment*, ya que no existe una traducción fiel para el término. Evidentemente, ello también representa un problema, ya que el término *evaluación* se utiliza de manera indistinta en las prácticas de los evaluadores profesionales al referirse a procesos propios de medición. Sin embargo, más allá del término utilizado para el fenómeno a gran escala en castellano (*large scale assessment* vs *evaluación a gran escala*), la evaluación a gran escala asume su base psicométrica y su ejercicio en pro de las decisiones que se quieran tomar con los resultados debido a la objetividad que con la elaboración de sus instrumentos pretende alcanzar.

3. La evaluación a gran escala y su relación con la rendición de cuentas

El uso de la evaluación a gran escala se ha implementado y mantenido por necesidades educativas y sociales de estandarización de contenidos y su verificación por medio del aprendizaje logrado por los alumnos, alentadas básicamente por decisiones de política gubernamental centrada en la rendición de cuentas, como a continuación se describe. Adicionalmente, la masificación de la educación ha traído consigo la necesidad de contar con elementos objetivos para la regulación del ingreso y egreso a las instituciones educativas.

La rendición de cuentas implica la obligación de reportar a otros, explicar o justificar sobre cómo se usan los recursos y los resultados que se obtienen con ellos. En el caso de la educación, la rendición de cuentas se relaciona con los resultados que de ella se generan y cómo se aprovechan con fines de una posible potencialización del mercado (Trow, 1996). La rendición de cuentas permite dar legitimidad a las instituciones, dándoles fuerza respecto a la forma en que manejan sus recursos y los resultados que se obtienen, incluso existe una tendencia en Europa respecto a enfocarse en la calidad (entendida como resultados) para la designación del presupuesto educativo, aunque de acuerdo con Trow (1996), la intención de relacionar la rendición de cuentas con la masificación de la educación superior se da en todo el mundo.

De acuerdo con Popham (2001, citado en Depascale, 2003), la evaluación a gran escala tiene la misión primaria de rendir cuentas, ello a través de la observación de los logros de los estudiantes como elemento de análisis del avance educativo. Esta facultad de la evaluación a gran escala de poner especial atención en los resultados logrados por los estudiantes favorece su implementación como un elemento de gobernanza, basado en el llamado *policy by number* (Saura y Luengo, 2015), en el cual el uso de mediciones permite establecer un monitoreo sobre las escuelas y el sistema educativo en general y así establecer acciones orientadas a cumplir las políticas educativas de importancia para el Estado.

De manera particular, la gobernanza debe entenderse como cualquier forma de gobierno que involucre a agentes públicos y privados, incluyendo las formas tradicionales de gobierno, así como aquella derivada de la imposición jerárquica de las medidas de información (Héritier, 2011; Khem, 2011). La gobernanza implica entonces la consideración de la participación de diversos agentes en la toma de decisiones sobre procesos importantes que anteriormente corrían únicamente a cargo del Estado (Khem, 2011). El Estado, entendido como el conjunto de instituciones dentro de un territorio y que ejerce poder sobre su población (Mann, 2007; O'Donnell, 2008) tradicionalmente se encarga de centralizar sus decisiones respecto a la política educativa de impacto a su población. En particular en América Latina, antes de la década de los ochenta del siglo pasado, los Estados se caracterizaban por ser proteccionistas y ofrecer una educación democrática (Vargas, 2002); sin embargo, como parte de las políticas económicas de dicha década, los gobiernos de dicha región se orientaron a disminuir el aparato estatal debido a los riesgos inflacionarios que las políticas proteccionistas podrían representar (Centro Latinoamericano de Administración para el Desarrollo, 2007). Como consecuencia, en la década de los noventa se desarrolló un proceso de integración económica internacional, que afectaría al ámbito educativo mediante la búsqueda de homogeneización de los sistemas educativos. Ello conllevaría a que la educación respondiera a los intereses del mercado.

El resultado se manifestó en que algunos elementos de gobernanza influirían en las políticas educativas en la región latinoamericana, aunque ello es un fenómeno que también se vive en otras regiones del mundo, tal es el caso de Europa. En los países de esta región también existe un proceso de estandarización de sistemas educativos, no como una imposición externa, sino como una acción voluntaria (Dale, 1999, citado en Bonal y Tarabini, 2013). Por ejemplo, en el caso particular de España, debido a su condición geográfica, ser parte de la región europea implica buscar compartir objetivos educativos para su población con otros países de la región, en particular siguiendo a aquellos países que han obtenido buenos resultados en evaluaciones internacionales (Bonal y Tarabini, 2013).

Las influencias internacionales como elemento relevante en el establecimiento de las políticas educativas han provocado que a nivel mundial se vivan cambios en la forma respecto en que se conceptualiza a la escuela pública y su financiamiento, así como la privatización escolar. Eso es a lo que Sahlberg (2011, citado en Saura y Luengo, 2015) denomina *Global Education Reform Movement*. El movimiento globalizado que se vive desde las últimas décadas del siglo pasado se basa en el libre mercado, es decir, en el factor competencia como desarrollo de mejores opciones para el consumidor, aspecto que no escapa a la educación, ya que de esa manera se fundamenta el discurso de una búsqueda constante de la calidad (Viñao, 2012). Estas prácticas neoliberalistas (Viñao, 2012) se basan en la idea de que lo privado se relaciona directamente con el bienestar social, mientras que lo público es algo ineficaz; como consecuencia, buscan la liberación y la desregulación de las transacciones económicas, en particular entre fronteras (Jessop, 2002). Ello se manifiesta en la educación mediante la reducción de políticas proteccionistas con el fin de promover una gestión basada en la competitividad que favorezca que los países se mantengan dentro de la dinámica económica global. En el caso de la administración pública en América Latina, bajo este esquema se comenzó a percibir a la educación, a la ciencia y a la tecnología como motores de desarrollo (Centro Latinoamericano de Administración para el Desarrollo, 2007).

La perspectiva neoliberal en educación defiende que los países de América Latina han tenido una deficiente capacidad administrativa y que el hecho de contar con sindicatos fuertes ha limitado la capacidad de rendir cuentas. En ese sentido, Winkler (2006) defiende que los Estados abiertos a la dinámica global deberán establecer objetivos mensurables, mejorar su monitoreo, así como diseñar esquemas de incentivos que los orienten hacia la obtención de financiamiento.

Bajo esta política comenzaron a establecerse mecanismos de monitoreo para que los Estados pudieran conocer los avances hacia los objetivos económicos deseados. De manera particular en lo que al ámbito educativo refiere, los sistemas educativos basados en la rendición de cuentas buscarían así el monitoreo de los logros alcanzados por la escuela, reflejados en los resultados de los estudiantes (Fuhrman y Elmore, 2004).

En este contexto, en América Latina se empieza a cuestionar el papel de la educación pública (en particular en educación superior) y empieza a ponerse especial atención al término de calidad en educación (Vargas, 2002). La calidad de la educación comienza a ser un factor importante, debido a la relevancia de asegurar la homogeneización de la educación superior en aras de alcanzar los estándares de formación que los procesos de globalización exigen. En este proceso de homogeneización o estandarización educativa, la evaluación juega un papel primordial, y es que la evaluación, al hacer uso de

estándares como elemento de juicio sobre el objeto evaluado¹, favorece que los sistemas educativos que se encuentran bajo la dinámica de rendición de cuentas puedan definir si los resultados educativos son suficientes o no respecto a lo esperado (Fuhrman y Elmore, 2004).

Las evaluaciones a gran escala permiten garantizar que el Estado tenga control y orden sobre el currículum escolar (Saura y Luengo, 2015), lo que facilita el cumplimiento de metas basadas en una economía globalizada. De manera particular, las evaluaciones de nivel internacional juegan un papel relevante en esta dinámica. La participación de los diversos países en evaluaciones internacionales se da por razones diversas, desde la búsqueda de fortalecer las políticas educativas nacionales, así como compararse con otros sistemas educativos o tener acceso a financiamiento internacional (Bonal y Tarabini, 2013). El desempeño de los estudiantes puede conocerse de manera fiable mediante instrumentos de evaluación de este tipo, por lo que se infiere que el avance escolar puede ser medido de forma justa mediante este tipo de herramientas (Fuhrman y Elmore, 2004); por supuesto, ello no excluye a las evaluaciones que se desarrollan de forma interna en cada país, ya que la rendición de cuentas puede realizarse de forma externa o interna, por lo que ambas implican una observación constante respecto al avance en alcanzar los resultados esperados (Trow, 1996). La evaluación en todos los niveles educativos, desarrollada tanto de forma externa como interna, facilitará entonces la consolidación de los sistemas educativos basados en la rendición de cuentas.

Actualmente, dentro de la sociedad se asume la relevancia de conocer el avance de resultados educativos, prueba de ello es, en el caso de México, el otorgamiento de la autonomía constitucional al Instituto Nacional para la Evaluación de la Educación (INEE) (Secretaría de Gobernación, 2013), lo que le da facultades para la evaluación del Sistema Educativo Nacional y con ello establecer políticas educativas. De esta manera, en México el Estado ha oficializado una preocupación centrada en que los alumnos alcancen un nivel mínimo de resultados educativos, la cual, como se ha comentado, se consolida mediante pruebas a gran escala. De acuerdo con autores como Díaz-Barriga (1994) y Villaseñor (2003), esta preocupación resulta una consecuencia que deriva de las políticas económicas neoliberales que, al menos en el caso de América Latina, comenzaron en los años ochenta.

Bajo una política neoliberal, la evaluación cumple una función de legalización de la regulación de la educación; es decir, es un reconocimiento administrativo al conocimiento (Díaz-Barriga, 1994), la idea es la de obtener un conocimiento *objetivo* del saber de cada uno de los estudiantes. Según Calero y Choi (2012), la evaluación a gran escala responde fuertemente a una política en la que se espera conocer los resultados de los insumos que los gobiernos destinan a la educación, siendo así un instrumento importante de la difusión y mantenimiento de políticas gubernamentales. Por lo tanto, el conocimiento de los resultados de los estudiantes será un indicador que favorezca la toma de decisiones respecto a la inversión que los Estados realizan directamente en sus sistemas educativos.

Como se comentó, algunos de los ejercicios más representativos de este tipo de evaluación son aquellos que se dan como una iniciativa internacional. En Latinoamérica

¹ Normalmente expresados en niveles de desempeño definidos dentro de las pruebas sistematizadas, o bien en puntajes mínimos esperados.

destacan los trabajos del Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE) de la Oficina Regional de la UNESCO para América Latina y el Caribe (OREALC), centrado en la evaluación de elementos en común de los currículos de los países latinoamericanos participantes en las áreas de Lenguaje, Matemáticas y Ciencias (Solano-Flores y Bonk, 2008). Los estudios que se han realizado en el LLECE constituyen el proyecto de evaluación internacional más importante hasta la fecha en Latinoamérica. Por su parte, la *Association for the Evaluation of Educational Achievement* (IEA) se desarrolló desde mediados de los años cincuenta en el Instituto de Educación de la UNESCO, en Hamburgo, hacia la posibilidad de un trabajo de evaluación a nivel internacional que permitiera obtener resultados estrictamente comparables, ello a través de la utilización de instrumentos equivalentes en contenido y grado de dificultad ante la insatisfacción por el uso de tasas de graduación o eficiencia terminal como único indicador de la calidad educativa. Con ese fin, en 1964 la IEA desarrolló el *First International Mathematics Study* (FIMS) (Saura y Luengo, 2015). En la actualidad, la IEA sigue desarrollando diversas evaluaciones a nivel internacional, como el *International Civic and Citizenship Education Study* (ICCS) (Schulz, Ainley, Fraillon, Kerr y Losito, 2010) o el *Trend in International Mathematics and Science Study* (TIMSS) (Ministerio de Educación de Chile, 2011). En ambos ejemplos hay participación de países pertenecientes a América Latina.

Por su parte, la Organización para la Cooperación y Desarrollo Económicos (OCDE) emprendió el proyecto denominado *Programme for the International Student Assessment* (PISA). Las pruebas se aplican cada tres años a estudiantes de 15 años desde el año 2000, centrando la atención en tres grandes áreas: Lectura, Matemáticas y Ciencias, con mayor énfasis en una de ellas en cada uno de los ciclos de aplicación² (INEE, 2013). La intención de la evaluación de PISA no se enfoca en los currículos de los países participantes, y es que es un programa centrado en conocer el avance de los jóvenes para hacer frente a las demandas de las sociedades de la información y del conocimiento. Independientemente del currículo del sistema educativo en cada país, se espera que los egresados del sistema educativo tengan herramientas que potencialicen su adaptación y éxito en un mundo globalizante (UNESCO, 2005), y PISA ha puesto énfasis en ello.

Los ejercicios internacionales son un indicador importante para la observación de los avances de los sistemas educativos, ya que con un estándar en común entre países participantes (ya sea conformado por temáticas coincidentes en los diversos currículos, como en el caso del LLECE, o bien por conceptos académicos como eje de comparación, como en el caso de PISA) se pueden realizar tendencias de los resultados de los estudiantes entre ciclos de aplicación y con ello analizar posibles tendencias de los avances que muestran los estudiantes, comparando los puntajes propios con contextos diferentes y similares al propio.

El monitoreo del avance educativo que se realiza mediante este tipo de evaluaciones es una práctica que permite conocer elementos genéricos de la educación para la toma de decisiones en materia de política educativa, dado que hace posible la identificación de avances o retrocesos en el aprendizaje de la población de un nivel educativo en

² Desde la aplicación de PISA 2012, en algunos países comenzó a evaluarse adicionalmente la Competencia financiera. En ciclos recientes se considerará también el área de Solución de problemas en colaboración como elemento de evaluación.

particular en una dimensión nacional o internacional, ya que hay ejemplos de pruebas nacionales en cada país que se usan con ese mismo fin, los cuales se abordan más adelante.

Al respecto, la reflexión que realiza Anderson (2005) permite comprender de mejor manera el interés de los ministerios de Educación en la evaluación a gran escala de corte nacional e internacional, al afirmar que en los sistemas basados en resultados reconocen públicamente el desempeño con el fin de otorgar recompensas (en algunos casos económicas) a las escuelas; de tal suerte que cuando los resultados de las evaluaciones otorgan beneficios para las escuelas o programas educativos, las evaluaciones comienzan a ser más solicitadas, incluso se pueden considerar como parte fundamental del sistema educativo (como en el caso del INEE en México), así como del proceso de toma de decisiones en materia de política educativa.

De esta manera, la evaluación (y en particular la evaluación a gran escala) representa una herramienta que ha permitido la consolidación de sistemas educativos basados en la rendición de cuentas, ya que favorecen la legitimación de una educación enfocada en la competencia económica. Por supuesto, bajo una política gubernamental globalizadora, ello responde a motivaciones de corte internacional. Como lo comenta Cepeda (2009) en su tesis doctoral, a consecuencia de la Declaración Mundial sobre Educación para Todos en el año de 1990, la UNESCO (a través de la Oficina Regional de Educación para América Latina y el Caribe, OREALC) desarrolló un discurso enfocado en evaluación, orientado particularmente en la definición de acciones de medición de resultados educativos que facilitaran la comparación entre escuelas e informar públicamente los resultados. De la misma manera, el Banco Mundial señaló la importancia de contar con mediciones para valorar la calidad del aprendizaje (Banco Mundial, 2000, citado en Cepeda, 2009). De esta forma se fomentaría que los países participaran en evaluaciones internacionales sobre desempeño educativo y se fomentara la creación de instituciones evaluadoras que ayudarán a consolidar una política de rendición de cuentas en función de los resultados obtenidos en educación. De igual modo, el Banco Interamericano de Desarrollo (2010) recomendó en 1996 la reestructuración de los sistemas educativos, argumentando la necesidad de que la formación de estudiantes se enfocara en el desarrollo de competencias entre escuelas, enfatizando así el discurso en la evaluación de la educación y sugiriendo orientar esta actividad hacia la participación de los países en proyectos internacionales con fines de obtener puntos de referencia para establecer metas nacionales de aprendizaje.

Claramente, en los casos mencionados, el fomento por la evaluación a gran escala se da por organismos que impulsan la apertura económica de los países en vías de desarrollo a los capitales extranjeros y, dado que la evaluación internacional fomenta la competencia económica, las sugerencias realizadas hacia los países en desarrollo favorecerán la apertura económica (Cepeda, 2009). Las políticas educativas no están entonces alejadas de las políticas económicas globalizadoras, por lo que en los discursos oficiales se busca que la enseñanza favorezca que los jóvenes sean competentes para hacer frente a un mercado mundial. Bajo esta mirada, en los años noventa se enfatizó el uso de pruebas a gran escala como el mejor indicador del avance de los sistemas educativos orientados a las competencias requeridas para una competitividad económica. A partir de esta visión es que esta práctica evaluativa se comenzó a considerar como una valiosa herramienta de monitoreo de las políticas educativas influenciadas por procesos de globalización económica.

En este sentido, la cultura mundializadora (Villaseñor, 2003) que tiene como eje la economía, permea a otras esferas de lo público y social, ya que es necesario que las instituciones (entre ellas las educativas) orienten sus esfuerzos a la obtención de los mayores beneficios económicos posibles. Al respecto, algunos de los proyectos de evaluación a gran escala (como los ya mencionados PISA y LLECE) se encuadran en la llamada sociedad del conocimiento, la cual, de acuerdo con Villaseñor (2003), se manifiesta también en la Economía del conocimiento. Es decir, el conocimiento que se genere se orienta a la producción de bienes y servicios, así como a su comercialización; de esta forma, el conocimiento se convierte en un elemento importante para la obtención de riqueza. Entonces, otro elemento de relevancia para la proliferación en el uso de la evaluación a gran escala también se da en este sentido, ya que es la herramienta que permite valorar si los sistemas educativos están cumpliendo con la función de preparar a los estudiantes para hacer frente a las demandas de competitividad que la economía del conocimiento implica.

La manera en que este tipo de evaluaciones influyen en acciones educativas en concreto puede verse reflejado en varios ejemplos. En España se han establecido algunos programas educativos como consecuencia de los resultados de PISA, como el Plan Nacional para Promover la Lectura (Bonafant y Tarabini, 2013). Por su parte, los análisis de Pereyra (2014), Tröhler (2009) y Meyer y Benavot (2013) dejan al descubierto la manera en que PISA ha sido un instrumento clave en la búsqueda de unificación de sistemas educativos bajo la mirada de políticas económicas neoliberales. PISA representa un medio de gobernación educativa de los sistemas educativos, estableciendo las naciones cuyos estudiantes se asemejan al ideal establecido por las naciones que controlan el 80% del PIB mundial (Pereyra, 2014). Mediante este instrumento a gran escala la OCDE ha logrado que los países que presentan la prueba orienten sus políticas educativas hacia donde la organización indica, debido a que los niveles de desempeño se desarrollan bajo un marco de competencia europeo. Dicho estándar se convierte entonces en el ideal educativo por alcanzar por parte de los ministerios de Educación.

Bajo esta lógica, la evaluación del aprendizaje a gran escala es una herramienta importante para las decisiones políticas relacionadas con la educación, ya que permite que el Estado valore la calidad de la educación a partir de los resultados de sus estudiantes en las pruebas aplicadas. En este sentido, la calidad irá en función de una estrategia educativa basada en la competencia, influenciada por políticas internacionales. Evidentemente, es un modelo de evaluación que favorece la implementación de acciones orientadas hacia la rendición de cuentas de un Estado, la cual establece como eje principal de competencia y calidad la observación del cumplimiento de estándares que el sistema educativo determina como adecuados.

4. Evaluación a gran escala en el sistema educativo mexicano

En el caso de México, la historia de la evaluación educativa de manera formal no es tan amplia como en otros países, como en Estados Unidos, por ejemplo. Sin embargo, los aspectos políticos que dieron pauta a su auge no distan mucho de lo ocurrido en otros contextos. Si bien desde la década de los sesenta en México comenzaron a utilizarse pruebas estandarizadas para la admisión de estudiantes a instituciones como la UNAM, la UIA y el ITESM (Martínez, 2008a), a partir de los años ochenta, a pesar de los

problemas económicos por los que atravesaba el país, México se dio a la tarea de enfrentar la masificación de la educación, así como la necesidad de atender las demandas de competitividad (Cepeda, 2009; Martínez, Cepeda, Fuentes y Burgos, 1995). Para explicar con mayor detalle la manera en que en México la evaluación (y en particular la evaluación a gran escala) comenzó a cobrar relevancia, vale recordar de forma general las políticas educativas del país en las últimas décadas.

Considerando la información contenida en la investigación desarrollada por Cepeda (2009), en México las políticas públicas en relación a la educación y que impactan directamente sobre los procesos de evaluación datan de los años setenta. A inicios de los años setenta ya se destinaba una gran proporción del Producto Interno Bruto del país a la educación, incluso comienza a plantearse la posibilidad de incrementar la educación básica hasta el nivel secundaria y el tema de la preocupación por la calidad aparece dentro del discurso público. Como parte de ese interés, en 1971 se crea el Instituto de Investigación Educativa, encargado de conformar estadísticas sobre la cobertura y reprobación en la educación básica, incluyendo también estudios explicativos sobre ese tipo de problemas. También se crea la Subsecretaría de Planeación Educativa dentro de la Secretaría de Educación Pública (SEP) y en 1974 aparece la Subdirección de Evaluación y Acreditación, área que comienza a desarrollar pruebas para el ingreso de alumnos a secundaria (aunque se habían iniciado desde 1972).

Ya en los años ochenta, debido a la crisis económica de principio de década en el gobierno de Miguel de la Madrid, el país comienza una importante apertura económica hacia otras naciones. Como consecuencia, hubo un debilitamiento de la industria local, un control de la inflación y la privatización de empresas públicas, así como una dependencia a instituciones crediticias, como el Fondo Monetario Internacional. El impacto de estas decisiones en términos educativos se manifestó en enfocar un interés hacia la calidad, eficiencia terminal y cobertura que el sistema educativo podía lograr. Desde ese momento comienza a valorarse la necesidad de un Sistema Nacional de Evaluación Educativa para sustentar mediante métodos objetivos la toma de decisiones en materia de política educativa enfocada a la calidad. Como consecuencia de esta intención, comenzaron a desarrollarse pruebas con fines de medición con diversos objetivos; por ejemplo, en el año 1984 se desarrolló el Instrumento de Diagnóstico y Clasificación para el Ingreso a la Educación Normal (IDCIEN).

Ya para el inicio de los años noventa se formalizó aún más el ingreso de empresas extranjeras a México a través del Tratado de Libre Comercio y, además, el país se incorporó como nación miembro a la OCDE. Con la finalidad de que el país se mantuviera competitivo frente a otros países en el marco de esos compromisos, se inició el Programa para la Modernización Educativa en 1992, enfocado en una reestructuración del sistema educativo y, a su vez, explicitando la necesidad de identificar las causas de la deserción educativa y el rezago escolar. En este marco se desarrolla el Diagnóstico para Alumnos de Nuevo Ingreso a la Educación Secundaria (IDANIS) y el programa de Carrera Magisterial, un programa de estímulos para docentes (Horbath y Gracia, 2014). También se desarrolla el Programa para Abatir el Rezago Educativo, realizando evaluaciones de aprendizajes en el aula, así como las Olimpiadas del Conocimiento Infantil, que comenzaron en 1993 y que implicaban ya una evaluación basada en el currículum. Entre los años 1995 y 2000 se realizó el Acuerdo Nacional para la Modernización de la Educación Básica, así como el Programa

Evaluación de la Educación Primaria en 1996, con la idea de generar estadísticas de tendencias educativas de 3° y 6° de Primaria.

De gran importancia fue la creación de la Ley General de Educación de 1993, la cual, bajo este contexto de competencia económica internacional, resalta la evaluación de la educación. Incluso, en el Plan Nacional de Desarrollo de la gestión de Carlos Salinas de Gortari, se señaló que el camino por seguir para determinar el financiamiento de las instituciones de educación superior sería a través de la evaluación. Como consecuencia se instauraron diversos organismos dedicados a la evaluación: por ejemplo, la Comisión Nacional de Evaluación (CONAEVA), los Comités Interinstitucionales para la Evaluación de la Educación Superior (CIEES) y el Centro Nacional de Evaluación para la Educación Superior (Ceneval) (Villaseñor, 2003). Adicional a los mecanismos de financiamiento que la evaluación educativa permitía establecer, uno de los mayores logros de esta política pública fue la implementación de un imaginario social de relación entre evaluación y financiamiento que posteriormente permitiría la asimilación de la sociedad de la evaluación como elemento fundamental de orientación hacia la mejora educativa.

Entre los años noventa y dos mil comenzó un especial interés en el desarrollo de instancias externas de evaluación de los resultados educativos de los estudiantes mexicanos. Con la creación del Ceneval se unificaron los procesos de selección de alumnos a la educación media superior de la zona metropolitana de la Ciudad de México a través de la Comisión Metropolitana de Instituciones Públicas de Educación Media Superior (COMIPEMS). Bajo la misma lógica, en el año 2002 se crea el INEE, que comenzó a desarrollar las actividades de evaluación que la SEP había llevado a cabo desde años antes (Horbath y Gracia, 2014), la cual, como se comentó, adquirirá posteriormente autonomía constitucional.

A partir del 2004 el INEE diseñó los Exámenes de Calidad y Logro Educativo (EXCALE) con el fin de informar sobre el avance del sistema educativo básico como conjunto, no respecto a cada estudiante en lo particular, así como identificar factores asociados al aprendizaje para brindar información para la toma de decisiones políticas. Por su parte, la SEP en el sexenio 2000-2006 desarrolló la Evaluación Nacional de Logro Académico en Centros Escolares (ENLACE). La prueba se desarrolló con el interés formativo de identificar el logro académico de alumnos de Primaria y Secundaria en algunas asignaturas clave con base en los planes y programas de la SEP, con el fin de atender las necesidades de cada alumno. Esta evaluación se consideró de forma censal con la finalidad de que los padres de familia y docentes tuvieran información directa sobre los logros alcanzados por los estudiantes, aprovechando así la experiencia que la Dirección General de Evaluación de Políticas de la SEP había alcanzado con evaluaciones desarrolladas desde los años setenta (SEP, 2007). Los objetivos primordiales de esta evaluación se centraban no solo en que los docentes y padres de familia tuvieran información sobre el avance de los logros educativos de sus hijos, sino que se desarrollaba con la intención de contribuir al desarrollo de mejores prácticas educativas desde el salón de clases, facilitando con esta información la toma de decisiones del proceso educativo y, con ello, “fortalecer la convergencia de esfuerzos nacionales y estatales a través del Sistema Nacional de Pruebas de Aprendizaje, de la Política Nacional de Evaluación y del Sistema Nacional de Evaluación” (SEP, 2007, p. 13). ENLACE se aplicó durante siete años seguidos en todo México a estudiantes tercero, cuarto, quinto y sexto de Primaria en las asignaturas: Español, Matemáticas y

Formación Cívica y Ética. También se aplicó en los tres años de Secundaria en las mismas asignaturas y cada año de aplicación se centraba en un área en particular.

La prueba ENLACE estuvo siempre sujeta a diversos comentarios por parte de la opinión pública: tanto positivos, considerándolo una medición relevante del estatus del logro de los estudiantes y necesaria para el conocimiento de la realidad educativa dentro de las escuelas del país, como negativos, debido a su uso enfocado únicamente como factor de reconocimiento social de las escuelas participantes. En enero del año 2014, el Secretario de Educación en turno, Emilio Chuayffett, anunció la cancelación de ENLACE, de manera que el INEE, ya con autonomía constitucional, sería el encargado de elaborar y aplicar las pruebas educativas en México para continuar con el monitoreo del logro de los estudiantes. Frente a ello, en el año 2015 el INEE sustituyó EXCALE y ENLACE por el Plan Nacional para la Evaluación de los Aprendizajes (Planea). Planea es una evaluación a gran escala enfocada en la educación obligatoria que pretende informar a la sociedad sobre el estado de la educación en función del logro de los estudiantes, así como brindar información a las autoridades educativas para el monitoreo y operación del sistema educativo de México y, por supuesto, contribuir al desarrollo de directrices sobre política educativa (INEE, 2016).

En México existe también el Ceneval, institución que desde 1994 comenzó, bajo la iniciativa de la Asociación de Universidades e Institutos de Educación Superior (ANUIES), a fungir como un agente externo de evaluación de la educación superior. Desde el año de su creación, el Ceneval desarrolla y califica el Examen Nacional de Ingreso a la Educación Media Superior (EXANI-I) a solicitud de la COMIPEMS (Hernández, Márquez y Palomar, 2006). Dicho instrumento permite desde hace más de 20 años tener un referente objetivo para la toma de decisiones respecto al ingreso a educación media superior en México, ya que actualmente se aplica en las instituciones públicas y privadas que así lo solicitan en todo el país. El EXANI también tiene dos versiones más, una centrada en el ingreso a educación superior (EXANI-II) y otra en estudios de posgrado (EXANI-III). Además del EXANI, el Ceneval desarrolla otro tipo de instrumentos a gran escala, incluso los elabora de acuerdo con las solicitudes propias de las instituciones educativas. Uno de los proyectos más grandes del centro es el Examen General para el Egreso de la Licenciatura (EGEL), el cual se centra en la medición de conocimiento y habilidades mínimos necesarios para el ejercicio profesional. De esta manera los EXANI favorecen la regulación del ingreso de estudiantes, mientras que los EGEL favorecen la del egreso de los programas educativos. En el proceso de consolidación de la evaluación a gran escala como herramienta de los sistemas educativos basados en competencia y en rendición de cuentas, los instrumentos desarrollados por el Ceneval favorecen la consolidación de una política de competitividad. De esta forma, la función de evaluación y acreditación de la calidad de la educación es un papel que el Estado canaliza a diversas agencias —en particular en educación superior—, fenómeno que se desarrolla en América Latina y el Caribe (en particular en Argentina, Chile, Brasil, Paraguay, Bolivia, Ecuador, Colombia, Costa Rica, El Salvador, México, Cuba, Trinidad y Tobago) (De la Garza, 2008).

Considerando las evaluaciones a gran escala internacionales, resalta la influencia del Banco Mundial y de la OCDE en programas estratégicos del sistema educativo en México. Por ejemplo, el Programa de desarrollo de la educación básica es un préstamo del Banco Mundial dirigido a apoyar la educación compensatoria en México, con el fin de buscar que la educación mexicana se acerque al ideal de los indicadores

internacionales establecidos por la OCDE. En particular, con este programa se ha apoyado al Consejo Nacional de Fomento Educativo (CONAFE) (Banco Mundial, 2005). Mediante esta acción el Banco Mundial afirma que está acercándose al cumplimiento de su objetivo de mejorar la calidad de la educación (entendida como la mejora del resultado de los estudiantes en evaluaciones estandarizadas) mediante los apoyos que brinda al Estado mexicano.

Por otra parte, mediante el financiamiento otorgado por el Banco Internacional de Reconstrucción y Fomento (BIRF, dependiente del Banco Mundial) se obtuvieron donaciones y becas y se apoyó a la evaluación de la educación que permitió una mejora en los resultados en la prueba PISA. Incluso, desde los años noventa, el Banco Mundial ha apoyado el Préstamo para Políticas de Desarrollo para la Educación Secundaria, con la finalidad de mejorar el acceso a la educación para mantener su competitividad (Banco Mundial, 2010). De esta forma, como lo comenta Jessop (2008, citado en Bonal y Tarabini, 2013), los tomadores de decisiones han hecho uso de PISA para construir nuevas formas de hegemonía. Además, este es un claro ejemplo de como la búsqueda de mejorar en los puntajes en evaluaciones se ha convertido en la principal meta de los cambios que los gobiernos realizan en materia educativa (Saura y Luengo, 2015).

De manera similar, la OCDE, como agente de gobernanza en la educación mexicana, ha emitido diversas recomendaciones que han facilitado la consolidación del uso de la evaluación a gran escala como medio de regulación educativa. De acuerdo con la OCDE (2010a), los elementos básicos que deben seguir los gobiernos son, entre otros aspectos, un sistema de rendición de cuentas basado en información objetiva de los resultados. Adicionalmente, se establece la necesidad de una reforma educativa en México basada en el hecho de que una gran proporción de estudiantes no finaliza la educación media superior, además de los bajos resultados mostrados en PISA. De la misma manera, dicho organismo sugiere continuar con el desarrollo de indicadores de resultados educativos nacionales (como ENLACE y EXCALE), ya que favorecen la atención a la mejora (OCDE, 2010b).

Este tipo de recomendaciones han facilitado la implementación de la evaluación a gran escala con fines de homogeneización de la educación. Según Anthony (2007), México debe trabajar para que sus estudiantes alcancen un nivel similar al de otros países de la OCDE, ya que la evaluación (junto con la autonomía de las escuelas y la rendición de cuentas de las mismas) ayudará a la mejora de la educación, propiciando así que la implementación de políticas sea más efectiva. Ello se ve reflejado en el desarrollo de programas tales como: Programa de Escuelas de Calidad (orientado a fortalecer la gestión escolar y el capital social de las escuelas), Programa Nacional de Lectura (que busca el fomento de la lectura y mejorar las prácticas pedagógicas relacionadas con ello) y el Programa Nacional de Formación Continua de los Maestros en Educación Básica (que pretende fortalecer la práctica docente). Estos se han desarrollado con la intención de mejorar el aprendizaje de los alumnos, aunque para ello es necesario establecer estándares de logro de los estudiantes y las escuelas como referente de su calidad (Miranda, 2007), lo que, como se comentó, son herramientas que favorecen la consolidación de un sistema educativo basado en la rendición de cuentas.

Previo a la Reforma Educativa mexicana que facilitó la autonomía del INEE en el año 2013, la OCDE (Santiago, McGregor, Nusche, Ravela y Toledo, 2012) sugirió continuar con el desarrollo de las evaluaciones nacionales como forma de comparación entre las calificaciones obtenidas por los estudiantes y los estándares nacionales que estas pruebas definen. También se sugirió a México continuar con el uso formativo de

evaluaciones a gran escala, con el fin de que los docentes reflexionen sobre los procedimientos necesarios para mejorar su enseñanza. En estas recomendaciones se da tal relevancia a la evaluación a gran escala que se sugiere evitar los enfoques normativos de evaluación (es decir, aquellos que se realizan dentro del aula por parte de los docentes) y centrarse en la relevancia de la evaluación por criterios, argumentando que algunos de los elementos clave para el éxito de una reforma educativa parten de la introducción exitosa de la alineación de los resultados de aprendizaje de los estudiantes y los estándares en el plan de estudios.

Considerando estos elementos históricos y de contexto global, la Reforma Educativa vivida en México en el año 2013 da un lugar preponderante a la evaluación de la educación. Esta tomó aún mayor fuerza que la mostrada en décadas previas, pues se considera que facilita la organización de todo el sistema educativo, así como el conocimiento del estatus *real* de la educación mexicana, para la implementación de políticas públicas basadas en evidencias o *policy by number*. De esta manera, los proyectos de evaluación a gran escala que se han generado desde la década de los ochenta y aquellos que se aplican en esta década representan una forma de control de la gestión del sistema educativo, en virtud de que es una herramienta que favorece el establecimiento de estándares unívocos, los cuales, a su vez, devienen de las políticas basadas en competencia consecuentes de la dinámica económica globalizadora. De esta forma, los resultados de las evaluaciones a gran escala fomentan esa determinación de un estándar o ideal de los resultados de aprendizaje de los jóvenes orientado hacia las dinámicas económicas globales, persiguiendo así una homogeneización de la formación de los jóvenes mexicanos, ya sea con fines de alcanzar estándares internacionales, o bien como medio de regulación interna del sistema educativo.

5. Conclusiones

Como se puede observar, el interés en evaluación como elemento al servicio del Estado mexicano se da a partir de la política educativa de dicho país a principios de los ochenta, ya que permitiría identificar problemas y jerarquizar las necesidades específicas de los ámbitos educativos con el fin de alcanzar la competitividad establecida por recomendaciones internacionales. Adicionalmente, desde estos años se vislumbraba ya la creciente demanda de educación como forma de movilidad social, la cual también podría ser regulada mediante la evaluación a gran escala, como sucede con el EXANI y el EGEL del Ceneval.

La evaluación a gran escala tiene entonces un papel crucial en el desarrollo de la política educativa mexicana, ya que asegura la verificación de que el aprendizaje que se pretendía enseñar mediante los programas educativos ha sido adquirido por los estudiantes, lo cual constituiría el elemento más palpable mediante el cual el sistema educativo podrá valorar sus propios avances hacia una sociedad globalizada. La evaluación a gran escala enfocada a los resultados de los estudiantes se convirtió en una herramienta efectiva de instrumentación de la cultura de la competencia enfocada a cumplir estándares internacionales. En México, además de los estudios como Planea del INEE, también hay participación en evaluaciones de corte internacional, tal es el caso de PISA o los estudios desarrollados por el LLECE, los cuales han tomado una importante fuerza política, dada su cualidad de analizar el avance del sistema educativo mexicano en términos de los resultados de aprendizaje de los estudiantes y su comparación con los de otras naciones.

De esta manera, la evaluación a gran escala en México ha tenido como fundamento de consolidación la intención de toma de decisiones, en particular como elemento de implementación de las políticas educativas orientadas a alcanzar estándares recomendados por instancias internacionales.

La consideración de la evaluación a gran escala como medio de implementación de políticas de competitividad y rendición de cuentas resulta inevitable, ya que las políticas educativas en torno a la evaluación permiten legitimar decisiones de gobernanza y ningún país ni su población puede mantenerse alejado de la dinámica económica globalizada. Frente a este panorama, vale repensar la utilidad de las pruebas a gran escala bajo un enfoque pedagógico centrado en lo formativo; es decir, valorar los resultados bajo una mirada crítica que tenga presente el papel de este tipo de evaluaciones. Resulta necesario considerar que la calidad de la educación no solo estará en función de factores exógenos (como los cataloga Márquez, 2004) que se encuadran únicamente en lo que socialmente se espera de la educación, en particular bajo un contexto globalizado. Como consecuencia, hay que recordar que la educación y su calidad no pueden referirse únicamente a los resultados de aprendizaje de los estudiantes bajo estándares unívocos (sobre todo en consideración de una sociedad tan desigual, como es el caso de la sociedad mexicana), sino también en función de la infraestructura escolar o el acceso y la permanencia en la escuela, entre otros elementos. Vale también considerar modelos de rendición de cuentas orientados hacia la responsabilidad que los sistemas educativos tienen sobre la población. En esa línea, Darling-Hammond y Snyder (2015) proponen un modelo orientado a facilitar el fomento a la responsabilidad que los sistemas educativos tienen sobre el aseguramiento de la educación que los niños merecen. Este tipo de propuestas pueden generar una alternativa que permita avanzar hacia la creación de sistemas alineados de evaluación y responsabilidad educativa. Resulta necesario considerar la posibilidad de reflexión sobre este tipo de alternativas, que ayudarán a valorar la evaluación a gran escala no solo como una herramienta para el control y la consolidación de políticas gubernamentales neoliberales.

Referencias

- Anderson, J. (2005). *La rendición de cuentas en la educación*. París: Instituto Internacional para la Planeación de la Educación.
- Anthony, H. (2007). Factores determinantes del aprendizaje y calidad de la educación en México. En F. Miranda, H. Anthony y A. López (Coord.), *Mejora de la calidad educativa en México. Posiciones y propuestas* (pp. 13-26). Ciudad de México: Consejo Mexicano de Investigación Educativa.
- Banco Interamericano de Desarrollo. (2010). *Reforma a la educación primaria y secundaria en América Latina y el Caribe*. Nueva York: BID.
- Banco Mundial. (2010). *México: Fortalecimiento del aporte de la educación a la economía del conocimiento*. Recuperado de <http://web.worldbank.org/WBSITE/EXTERNAL/BANCOMUNDIAL/NEWSSPANISH/0,,contentMDK:22899181~menuPK:3327642~pagePK:34370~piPK:34424~theSitePK:1074568,00.html>
- Bonal, X. y Tarabini, A. (2013). The role of PISA on shaping hegemonic educational discourses, policies and practices: The case of Spain. *Research in Comparative and International Education*, 8(3), 335-341. doi:10.2304/rcie.2013.8.3.335.

- Calero, J. y Choi, A. (2012). La evaluación como instrumento de política educativa. *Presupuesto y Gasto Público*, 67, 29-42.
- Centro Latinoamericano de Administración para el Desarrollo. (2007). Una nueva gestión pública para América Latina. En C. H. Acuña (Comp.), *Lecturas sobre el Estado y las políticas públicas: Retomando el debate de ayer para fortalecer el actual* (pp. 403-424). Buenos Aires: Jefatura del Gabinete de Ministros.
- Cepeda, B. (2009). *La política de evaluación externa de los aprendizajes en educación básica. Una aproximación a la comprensión de su racionalidad. Periodo 2000-2006* (Tesis doctoral no publicada). Universidad Nacional Autónoma de México, Ciudad de México.
- Darling-Hammond, L. y Snyder, J. (2015). Meaningful learning in a new paradigm for educational accountability: An introduction. *Education Policy Analysis Archives*, 23(7). doi:10.14507/epaa.v23.1982
- DePascale, C. (2003). The ideal role of large-scale testing in a comprehensive assessment system. *Association of Test Publishers*, 5(1), 1-11.
- De la Garza, J. (2008). Evaluación y acreditación de la educación superior en América Latina y el Caribe. En C. Tünnermann (Ed.), *La educación superior en América Latina y el Caribe: Diez años después de la Conferencia Mundial de 1998* (pp. 175-222). Cali: Sello Editorial Javeriano.
- Díaz Barriga, A. (1994). Una polémica en relación al examen. *Revista Iberoamericana de Educación*, 5, 161-181.
- Fuhrman, S. y Elmore, R. (Eds.). (2004). *Redesigning accountability systems for education*. Nueva York: Teachers College Press.
- Héritier, A. (2011). New modes of governance in Europe: Policy-making without legislating? *Preprints aus der Max-Planck-Projektgruppe Recht der Gemeinschaftsgüter*, 14.
- Hernández, J., Márquez, A. y Palomar, J. (2006). Factores asociados con el desempeño académico en el EXANI-I. *Revista Mexicana de Investigación Educativa*, 11(29), 547-581.
- Horbath, J. y Gracia, M. (2014). La evaluación educativa en México. *Revista de Relaciones Internacionales, Estrategia y Seguridad*, 9(1), 59-85.
- Instituto Nacional para la Evaluación de la Educación. (2013). *México en PISA 2012*. Ciudad de México: Autor.
- Instituto Nacional para la Evaluación de la Educación. (2016). *Planea: Una nueva generación de pruebas. ¿Cómo y cuándo se evalúa?* Recuperado de http://www.inee.edu.mx/images/stories/2015/planea/fasciculosnov/Planea_2.pdf
- Jessop, B. (2002). Liberalism, neoliberalism, and urban governance: A state-theoretical perspective. *Antipode*, 34(3), 452-472. doi:10.1111/1467-8330.00250.
- Khem, B. M. (2011). Gobernanza, ¿qué es? ¿Es importante? En B. M. Khem (Ed.), *La gobernanza en la enseñanza superior* (pp. 21-25). Barcelona: Universitat de Barcelona.
- Mann, M. (2007). El poder autónomo del Estado: Sus orígenes, mecanismos y resultados. En M. Mann (Ed.), *Lecturas sobre el Estado y las políticas públicas* (pp. 55-78). Buenos Aires: Jefatura del Gabinete de Ministros.
- Martínez, F. (2001). Evaluación educativa y pruebas estandarizadas. Elementos para enriquecer el debate. *Revista de la Educación Superior*, 30(4), 71-85.
- Martínez, F. (2008a). *Evaluación a gran escala y evaluación en el aula. Los usos de Excale para la mejora educativa*. Recuperado de <http://www.inee.edu.mx/explorador/propuesta2008/Textos/EvaluacionGranEscala.pdf>

- Martínez, F. (2008b). *La evaluación de aprendizajes en América Latina*. Ciudad de México: Instituto Nacional para la Evaluación de la Educación.
- Martínez, F. y Santos, A. (2009). Evaluación e investigación educativa. En A. Alba y R. Glazman (Coord.), *¿Qué dice la investigación educativa?* (pp. 265-304). Ciudad de México: Consejo Mexicano de Investigación Educativa.
- Martínez, F., Cepeda, B., Fuentes, G. y Burgos, R. (1995). Evaluación del aprendizaje. En M. Rueda (Coord.), *La investigación educativa en los ochenta, perspectivas para los noventa. Procesos de enseñanza y aprendizaje I* (pp. 11-19). Ciudad de México: Consejo Mexicano de Investigación Educativa.
- Márquez, A. (2004). Calidad de la educación superior en México. ¿Es posible un sistema eficaz, eficiente y equitativo? Las políticas de financiamiento de la educación superior en los noventa. *Revista Mexicana de Investigación Educativa*, 9(21), 477-500.
- Meyer, H. y Benavot, A. (2013). *PISA, power and policy: The appear of global governance of education*. Oxford: Symposuim Books.
- Ministerio de Educación de Chile. (2011). *Resultados de TIMSS 2011. Chile. Estudio Internacional de Tendencias en Matemática y Ciencia*. Santiago de Chile: Autor.
- Miranda, F. (2007). Mejora del aprendizaje y la calidad de la educación: Una visión política educativa. En F. Miranda, H. Anthony y A. López (Coord.), *Mejora de la calidad educativa en México. Posiciones y propuestas* (pp. 119-138). Ciudad de México: Consejo Mexicano de Investigación Educativa.
- O'Donnell, G. (2008). Algunas reflexiones acerca de la democracia, el Estado y sus múltiples caras. *Revista del CLAD Reforma y Democracia*, 42, 5-30.
- Organización para la Cooperación y el Desarrollo Económicos. (2010a). *Acuerdo de cooperación México-OCDE para mejorar la educación de las escuelas mexicanas. Resúmenes ejecutivos*. París: Autor.
- Organización para la Cooperación y el Desarrollo Económicos. (2010b). *Mejorar las escuelas. Estrategias para la acción en México. Resumen ejecutivo*. París: Autor.
- Patrinos, H. A., Shapiro, J. y Moreno Trevino, J. (2005). *Educación compensatoria para estudiantes desfavorecidos: Pruebas de un estudio de evaluación de impacto llevado a cabo en México*. Washington D. C.: Banco Mundial.
- Pereyra, M. (2014, septiembre). PISA a examen: ¿Evento mediático, nueva forma de gobernanza, o propuesta de cambio escolar? Comunicación presentada en el *Congreso Internacional de Educación*. Evaluación. Universidad Autónoma de Tlaxcala, Tlaxcala.
- Ravela, P. (2006). *Para comprender las evaluaciones educativas. Fichas didácticas. Programa de promoción de la Reforma Educativa en América Latina y el Caribe*. Recuperado de http://www.iae.org.mx/2012/Dip1DGIFA/nov1/S102_Ravela_Pedro_Fichas_didacticas.pdf
- Santiago, P., McGregor, I., Nusche, D., Ravela, P. y Toledo, D. (2012). *Revisiones de la OCDE sobre evaluación en la educación*. Ciudad de México: INEE.
- Saura, G. y Luengo, J. (2015). Política global más allá de lo nacional. Reforma educativa (LOMCE) y el régimen de estandarización (OECD). *Bordón*, 67(1), 135-148. doi:10.13042/bordon.2015.67109.
- Schulz, W., Ainley, J., Fraillon, J., Kerr, D. y Losito, B. (2010). *ICCS 2009 International report: Civic knowledge, attitudes, and engagement among lower-secondary school students in 38 countries*. Amsterdam: IEA.

- Secretaría de Gobernación. (2013). DECRETO por el que se reforman los artículos 3o. en sus fracciones III, VII y VIII; y 73, fracción XXV, y se adiciona un párrafo tercero, un inciso d) al párrafo segundo de la fracción II y una fracción IX al artículo 3o. de la Constitución Política de los Estados Unidos Mexicanos. Ciudad de México: Diario Oficial de la Federación.
- Secretaría de Educación Pública. (2007). *Manual técnico ENLACE*. Ciudad de México: Autor.
- Solano-Flores, G. y Bonk, W. (2008). *Evaluation of the Latin American Laboratory for the Evaluation of Educational Quality (LLECE)*. Santiago de Chile: OREAC/UNESCO.
- Tröhler, D. (2009). Armonizar el globo educativo. Política mundial, rasgos culturales y los desafíos de la investigación educativa. *Profesorado. Revista de Currículum y Formación del Profesorado*, 13(2), 1-14.
- Trow, M. (1996). Trust, markets and accountability in higher education: A comparative perspective. *Higher Education Policy*, 9(4), 309-324. doi:10.1016/s0952-8733(96)00029-3.
- Vargas, A. (2002). La acreditación: Una forma de estandarizar la educación. *Revista Educación*, 26(2), 245-254. doi:10.15517/revedu.v26i2.2923.
- Villaseñor, G. (2003). La evaluación de la educación superior: Su función social. *Reencuentro*, 36, 20-29.
- Viñao, A. (2012). El desmantelamiento del derecho a la educación: Discursos y estrategias neoconservadoras. *Revista Internacional de Ciencias Sociales*, 31, 97-107.
- Winkler, D. (2006). Fortalecer accountability (rendición de cuentas) en la educación pública. En J. Corvalán y R. McMeekin (Eds.), *Accountability educacional: Posibilidades y desafíos para América Latina a partir de la experiencia internacional* (pp. 99-123). Santiago de Chile: PREAL.

Problemas Éticos y Representaciones Sociales de la Evaluación en la Universidad

Ethical Problems and Social Representations of Evaluation at University

Nelly Patricia Bautista Cárdenas*

Universidad Cooperativa de Colombia

Este artículo es producto de una investigación sobre la evaluación en educación superior desde el aspecto cultural a partir de las representaciones sociales. La motivación para la reflexión fue la sobrevaloración de la calificación frente al aprendizaje percibida en la población estudiantil. La primacía en la nota conduce a crear un vacío en el campo ético y cívico además de anular la función formativa de la evaluación en el curso del desarrollo de una carrera universitaria. Por eso, el objetivo de la investigación consistió en evidenciar la representación social de la evaluación como fin y no como medio en el contexto universitario, permitiendo así exhortar a las entidades universitarias a que promuevan acciones que neutralicen las conductas violentas o delictivas de quienes pretenden obtener calificaciones aprobatorias o certificaciones por encima de la lógica educativa y de los derechos humanos. La metodología utilizada es de orden cualitativo en el marco de las representaciones sociales, lo cual permitió relacionar los discursos constitucionales y las normas institucionales con el sentido común que construyen docentes y estudiantes ante el proceso evaluativo, encontrándose elementos subyacentes de poder, manipulación, temor, plagio y conceptos de la evaluación diferentes al que predicán la teoría y la norma.

Palabras clave: Ética, Evaluación universitaria, Derechos humanos, Metodología de la evaluación, Representaciones sociales.

This article is the result of research about Evaluation in higher education from the cultural aspect seen from the Social Representations. The motivation for reflection was overvalued about the perceived qualifying in the student population. The primacy in note leads to create a vacuum in the ethical and civic field and will void the formative role of evaluation in the course of development of college. Therefore, the purpose of investigation is to show the social representation of the evaluation as an end and not a means, in the university context allowing institutions to urge the university to promote actions to neutralize violent or criminal behavior of those that seek passing grades or certifications over educational logic and human rights. The methodology used is qualitative in the framework of social representations, allowing relate the constitutional discourse and institutional rules with common sense that build teachers and students to the evaluation process, finding underlying elements of power, manipulation, fear, plagiarism and concepts of the different assessment to preaching theory and policy.

Keywords: Ethics, University evaluation, Human rights, Evaluation Methodology, Social representations.

1. Introducción

La evaluación forma parte de la historia del ser humano desde su nacimiento, lo que conduce a construcciones sociales que han hecho los actores de la educación en forma paralela a los enfoques teóricos y a las normas que le subyacen, reconociendo y moldeando las representaciones sociales que evidencia el mundo de la vida, complejo, contradictorio y subjetivo, que moldea el mundo normativo de la sociedad.

En este sentido, la investigación evidenció el aspecto ético de la evaluación en el contexto universitario con base en datos recogidos en la labor docente, donde se encontró que algunos estudiantes del programa de psicología de la Universidad Cooperativa de Colombia, al dar prioridad a la calificación sobre el aprendizaje y al Diploma sobre el profesionalismo, trasgreden los códigos éticos, morales y sociales que subyacen no solo a las normas institucionales sino también a los derechos humanos y los principios de dignidad humana, en cuyos cimientos filosóficos se sustenta el concepto de autonomía, derechos y libertades contenidos en la Carta Magna colombiana, cuyo disfrute implica la responsabilidad sobre los alcances de las acciones individuales y colectivas.

Los académicos entienden que la evaluación es la que diagnostica qué, cómo, por qué y cuándo enseñar, lo cual le da un lugar destacado a este acto formativo. En palabras de Arturo de la Orden (1981), “la evaluación, al prescribir realmente los objetivos de la educación, determina, en gran medida lo que los alumnos aprenden y cómo lo aprenden, lo que los profesores enseñan y cómo lo enseñan, los contenidos y los métodos” (p. 120). Es decir, la actividad educativa está en alto grado canalizada por la evaluación, lo que nos hace un llamado al aspecto ético que le subyace porque, al ser un elemento mediador en la educación, repercutirá en la formación integral del futuro profesional.

Se construyen normas generales desde el Ministerio de Educación e institucionales en cada centro educativo para delimitar parámetros claros y definidos en la labor educativa; se propende por una evaluación orientada al aprendizaje que se constituya en sí misma como un momento y una oportunidad para aprender. No obstante, el ser humano pone resistencia a los parámetros impuestos y complejiza su vida mediante la construcción de significados, símbolos y discursos que se entremezclan con las disposiciones oficiales, dando lugar a unas prácticas culturales que se hacen difíciles de reconocer en las acciones de los individuos. Es ahí donde ubicamos la tradición cultural expresada en las representaciones sociales que se entretajan con evaluación en la educación superior para reconocer el universo infinito que se instala en la subjetividad de los actores de la educación.

2. Marco teórico

2.1. Concepto de evaluación de los aprendizajes

La evaluación de los aprendizajes es un proceso complejo en el que intervienen elementos y factores variados como la institución, los profesores, los estudiantes, la enseñanza, los aprendizajes y, en cada uno de ellos, todos los significados que se conjugan y que abre una brecha entre significados culturales y su concepto teórico.

Según Correa, Puerta y Restrepo (2002) “el término *evaluación* es una palabra elástica que se emplea para abarcar muchas clases de juicios” (p. 27). Se habla de evaluación de

la ejecución de tareas, evaluación de la calidad de la educación, evaluación del clima organizacional. En ese sentido, por evaluación entienden algunos autores el control del éxito o fracaso de programas, proyectos o instituciones, para identificar y medir resultados. “Así, la evaluación es un esfuerzo por reconocer qué cambios se presentan durante y después de un programa de acción y qué parte de dichos cambios pueden atribuirse al programa” (Correa et al., 2002, p. 27).

A continuación, se presentan algunas de las definiciones que dan los autores más destacados que evidencian el cambio que ha tenido este tema en las últimas décadas.

Arturo de la Orden (1981): “Proceso de recogida y análisis de información relevante para describir cualquier faceta de la realidad educativa y formular un juicio sobre su adecuación a un patrón o criterio previamente establecido como base para la toma de decisiones” (citado por Chavarría y Borrel, 2003, p. 52).

Pedro Lafourcade (1984): “Etapa del proceso educacional que tiene por fin comprobar de modo sistemático en qué medida se han logrado los resultados previstos en los objetivos que se hubieren especificado con antelación” (citado por Ruiz, 2014, p. 2).

Robert Stake y Terry Denny definen el concepto de evaluación como el “descubrimiento de la naturaleza y el valor de algo”, “una ayuda para descubrir el mérito y las deficiencias de un programa” (citado por Proppe, 1990, p. 325).

José Gimeno Sacristán (1996):

La evaluación hace referencia a cualquier proceso por medio del cual alguna o varias características de un alumno, de un grupo de estudiantes, de un ambiente educativo, de objetivos educativos, de materiales, profesores, programas, etc., reciben la atención del que evalúa, se analizan y se valoran sus características y condiciones en función de unos criterios o puntos de referencia para emitir un juicio que sea relevante para la educación.
(p. 338)

Miguel Ángel Santos Guerra (1993, p. 6): “Proceso de diálogo, comprensión y mejora, como discusión y reflexión compartida de todos los que están implicados en la actividad evaluada”.

Juan M. Álvarez (2001): La evaluación es una oportunidad para que quienes aprenden pongan en práctica “sus conocimientos y se sientan en la necesidad de defender sus ideas, sus razones, sus saberes”. También es el momento en el que “afloren las dudas, las inseguridades, las ignorancias, si realmente hay intención de superarlas” (p. 13).

María José Rodríguez Conde (2005): “Conjunto de procesos sistemáticos de recogida, análisis e interpretación de información válida y fiable, que en comparación con una referencia o criterio nos permita llegar a una decisión que favorezca a la mejora del objeto evaluado” (citado por Ruiz, 2014, p. 2).

Mario de Miguel Díaz (2006): “Proceso planificado, integral y pertinente a las competencias que se desean alcanzar. Se desarrolla a través del planteamiento de tareas o desafíos que el estudiante debe resolver, necesitando para ello un conjunto integrado de conocimientos, destrezas y actitudes” (citado por Ruiz, 2014, p. 2).

Elena Cano (2008): “Proceso que utiliza diversidad de instrumentos e implica a diferentes agentes, con el propósito de proporcionar información sobre la progresión en el desarrollo de la competencia y sugerir caminos de mejora” (citado por Ruiz, 2014, p. 2).

Tiburcio Moreno Olivos (2009): “La evaluación es el medio por el cual valoramos y conocemos una situación educativa, bien en su proceso, bien en un momento determinado, pudiendo conocer la efectividad de la actividad” (pp. 564-565).

Como se observa con las anteriores definiciones, el concepto de evaluación ha cambiado a través de los años. Inicialmente se entendía como medición del evaluado para modificarse progressively hacia la idea de una evaluación que mide a todos los implicados en esta actividad y cuya finalidad es servir de referencia para tomar decisiones y acciones de mejora.

2.2. Estrategias de la evaluación

Según Ronald Feo (2010), para el diseño de la estrategia de evaluación el profesor emplea de manera básica diferentes instrumentos que permitan recabar información. Esta información se puede alcanzar a través de técnicas evaluativas que conlleven la observación y análisis de los procesos o productos finales elaborados en las actividades. Es necesario resaltar que la estrategia de evaluación gira en torno a criterios evaluativos congruentes con la actividad evaluativa diseñada. Se utiliza una gran variedad de técnicas e instrumentos evaluativos como: observación, portafolios, mentefactos, entrevistas, encuestas, informes, estudios de caso, tareas, ensayos, exposiciones y pruebas escritas.

2.3. Funciones de la evaluación

La evaluación de los aprendizajes cumple diferentes funciones, entre ellas encontramos las definidas por Gimeno Sacristán (1996, p. 364):

Definición de los significados pedagógicos y sociales; funciones sociales; poder de control; funciones pedagógicas; creadora del ambiente escolar; diagnóstico; recurso para la individualización; afianzamiento del aprendizaje; función orientadora; base de pronósticos; ponderación del currículum y socialización profesional; funciones en la organización escolar; proyección psicológica; y apoyo de la investigación.

A continuación se explican estas funciones Gimeno Sacristán, 1988,:

- ✓ *Definición de los significados pedagógicos y sociales:* “Los usos y resultados de la evaluación [...] generan realidad y sirven para pensar, hablar, investigar, planificar y hacer política educativa sobre la educación. La evaluación dota de contenido a la idea de ‘excelencia escolar’”, que se utiliza para hablar de lo que resulta mejor o peor (pp. 364-365).
- ✓ *Funciones sociales:* En una sociedad en la que “la acreditación del saber y de las formas de ser o comportarse, expresan la posesión de un capital cultural y de valores que se cotizan en la sociedad”, los resultados de evaluación de los aprendizajes repercutirán en las “oportunidades de entrada al mercado laboral” ya que mide los niveles de competencia del evaluado (p. 365).
- ✓ *Poder de control:* Se evidencia en “la actitud de algunos profesores al imponerlas, por la forma de realizarlas, la potestad de corregir respuestas interpretables” (p. 366), la posibilidad de discutir los resultados, se pueden convertir en un instrumento de dominación. El evaluador manifiesta los criterios sobre lo que es correcto o no.
- ✓ *Funciones pedagógicas:* “La evaluación cumple un papel de regulación de los procesos de enseñanza y de aprendizaje, es decir, de reconocimiento de los cambios que deben introducirse en este proceso a fin de que cada estudiante aprenda de forma significativa” (p. 367).

- ✓ *Creación del ambiente escolar*: Es “la proyección que tiene la evaluación en las relaciones interpersonales” entre profesores y alumnos porque sus resultados sirven de referencia en la estructuración de las relaciones sociales. “Las normas y valores por las que funciona el grupo escolar tienen mucho que ver con el clima de evaluación” (p. 367).
- ✓ *Diagnóstico*: Esta función es la que da a conocer el progreso y “el funcionamiento de los procesos de aprendizaje con el fin de intervenir en su mejora”. Se utiliza en el sentido de conocer al estudiante, sus saberes previos, sus condiciones personales, familiares y sociales; para obtener una perspectiva global de las personas en su propio contexto; conocimiento del curso para corregir y mejorar los procesos; proyección del estado final del alumno; y diagnóstico de las cualidades de los alumnos para la conformación de grupos (p. 367).
- ✓ *Recurso para la individualización*: “La función diagnóstica de la evaluación permite la adaptación de la enseñanza a las condiciones del alumno y a su ritmo de progreso así como el tratamiento de dificultades particulares” (p. 369).
- ✓ *Afianzamiento del aprendizaje*: La evaluación se encuentra dentro del proceso de seguir aprendiendo: Cuando la evaluación trata una lección concreta, activa en el estudiante su atención sobre aspectos importantes del contenido; si la evaluación es sobre módulos o cursos completos, activa la motivación del alumno hacia la asignatura. A largo plazo, se estimula la habilidad de retener y aplicar en contextos y formas diversas el material aprendido.
- ✓ *Función orientadora*: Al detectar las cualidades y competencias en ciertas asignaturas, la evaluación sirve de guía para tomar decisiones apropiadas a la elección de tipos de estudios o la creación de materias optativas.
- ✓ *Base de pronósticos*: Los resultados de las evaluaciones son referentes para crear imágenes de los alumnos en el medio escolar y en el exterior; sobre su valía académica o su ‘excelencia’. Estas informaciones generan expectativas que condicionan la valoración de datos posteriores.
- ✓ *Ponderación del currículum y socialización profesional*: Las maneras de evaluar ponen de manifiesto los esquemas implícitos, la idea sobre conocimiento relevante y las ideas que se tienen sobre el “aprendizaje ideal”. Estas reflexiones actúan como filtros en el desarrollo del currículum (p. 371).
- ✓ *Funciones en la organización escolar*: La formación de “muchos alumnos durante un largo período de sus vidas implica enfrentarse a una heterogeneidad importante de condiciones, capacidades y posibilidades queriendo transmitirles aprendizajes muy diversos, reclama organizar alumnos y conocimientos”, y es allí donde la evaluación refleja la complejidad de la organización del programa (p. 374).
- ✓ *Proyección psicológica*: “La evaluación tiene repercusiones psicológicas en la motivación, en la modelación del autoconcepto personal, en las actitudes del alumno, en la generación de ansiedad y en la acentuación de conflictos o rasgos patológicos” (p. 376).
- ✓ *Apoyo de la investigación*: Los resultados de la enseñanza y la evaluación son una fuente de información y conocimiento sobre cómo funciona el sistema educativo, siendo necesario considerar todo tipo de datos tanto cuantitativos como cualitativos para que realmente sean un apoyo fidedigno a la investigación (p. 378).

2.4. Ética en la Evaluación

Introducir el tema de la ética en la acción evaluadora es abordar la moral como costumbre y norma instituida en el seno de la sociedad, con el fin de ubicar lo que desde la cultura se asume como vital para defender la óptima formación integral del individuo y del grupo de estudiantes dentro una sana convivencia. La evaluación entonces se torna un pretexto para discriminar entre lo bueno y lo malo, lo idóneo y lo insuficiente, lo legal y lo ilegal, por lo que se armoniza con la norma institucional y la jurídica ya que ambas buscan regular las acciones e interacciones de los actores sociales, que en este caso señala a quienes conforman la comunidad universitaria.

El ejercicio evaluativo se rige por tendencias nacionales e internacionales trascendiendo el acto inicial de promover a los estudiantes. En este sentido, desde el ámbito nacional la evaluación está atravesada por determinadas posturas políticas que exhiben su intervención en la normativa y en la práctica académica. Las tendencias internacionales, por su parte, fundamentan la formulación de dichas políticas imponiendo parámetros de competitividad a los profesionales, en el aspecto del saber-hacer. En así que el fenómeno evaluativo se ve permeado por los discursos éticos y políticos que definen sus funciones.

Los discursos políticos tienen profundo impacto en la representación de la función ideal del proceso evaluativo y se desprenden de la misión y visión institucional, así como de estudios profundos en el tema. No obstante, es común encontrar que muchas veces los agentes evaluadores los infravaloran, centrándose fundamentalmente en lo técnico como forma de comprobar el saber profesional. Así se da prioridad al saber objetivo sobre el pensamiento ético, creativo y crítico.

Esta forma de entender la medición del conocimiento tiene una profunda carga ideológica y de poder que direcciona la manera como se seleccionan las técnicas, las herramientas, las formas y los procedimientos para desarrollar la evaluación, así como la valoración que se haga de los procesos y de los resultados. La evaluación de los aprendizajes no puede reducirse a la aplicación de instrumentos con el fin de recoger información, ya que esto coarta la responsabilidad ética y dificulta la formalización de metas claras que provean de significado las acciones educativas, restringiendo también la toma de decisiones.

Cuando el ejercicio evaluativo se mueve de manera sólida, apoyado por presupuestos teóricos y críticos que explicitan su sentido, se convierte en el seguimiento adecuado para dar cuerpo a la función formativa de la evaluación y dispone la apertura al pensamiento ético que redunde en compromiso profesional y social. No obstante, como lo demuestra la teoría y el método de las representaciones sociales (Jodelet, 1984), “el mundo no obra de acuerdo a los parámetros establecidos, ni a la razón sino a un complejo entramado que es lo que llamamos el mundo de la vida” (Bautista, 2015). La evaluación, como algo que está presente en la experiencia del ser humano desde el inicio de su existencia, en el contexto universitario también está cargada de sentido y no se limita únicamente a la selección, clasificación y orientación al grupo de estudiantes, sino que se ve permeada por acciones que en ocasiones se distancian del objetivo formativo, pasando a ser una mezcla de razón y sin razón evidenciada en acciones y emociones alejadas del ideal.

Es de recordar también en el aspecto ético que la academia reproduce y legitima las desigualdades sociales y, siendo la evaluación una herramienta que permite ordenar y clasificar a los estudiantes, se convierte en un parámetro decisivo de su futuro y su

posición social (Murillo e Hidalgo, 2015). Por eso, como lo plantean estos investigadores, se debe seguir un enfoque de evaluación para justicia social que busque recuperar el valor ético del ejercicio evaluativo desde todos los actores de la educación.

4. Representaciones sociales

El concepto de representaciones sociales se sustenta teóricamente en tres corrientes básicas: la *etnopsicología de Wundt*, el *interaccionismo simbólico de Mead* y las *representaciones colectivas de Emile Durkheim* (Mora, 2002). La etnopsicología o psicología de los pueblos explica la evolución de la mente del hombre, la importancia del lenguaje, el pensamiento y la comunicación. Por su parte, la contribución del interaccionismo simbólico de George Mead fue su tesis de la intersubjetividad basada en lo que llamó la conversación interior desde tres interlocutores permanentes que eran el yo, el mí y el otro generalizado o realidad social, resaltándose el valor de la simbología en la comunicación. Por último, están los aportes de Emile Durkheim desde su teoría de las representaciones colectivas, donde sostiene que lo colectivo no puede ser reducido a lo individual imponiéndose parámetros sociales de pensamiento y acción en el individuo desde su nacimiento (Mora, 2002).

Estas influencias marcaron el desarrollo de la teoría que propuso Serge Moscovici sobre representaciones sociales. Para Moscovici (1979), estas son generadas en procesos de intercambios e interacciones, por lo cual decide desligarlas de la categoría preestablecida y estática que traía la visión clásica de Durkheim. Así, Moscovici sintetiza la construcción de representaciones sociales en los siguientes postulados (Araya, 2002, p. 26):

- a) “Nuestro conocimiento, más que ser reproductor de algo preexistente, es producido de forma inmanente en relación con los objetos sociales que conocemos”.
- b) “La naturaleza de esa generación y construcción es social, esto es, que pasa por la comunicación y la interacción entre individuos, grupos e instituciones”.
- c) El lenguaje y la comunicación son “mecanismos en los que se transmite y crea la realidad, por una parte, y [...] marco en que la realidad adquiere sentido, por otra”.

Con base en las anteriores premisas, Moscovici definió las representaciones sociales como una modalidad particular del conocimiento cuya función es la elaboración de los comportamientos y la comunicación entre los individuos. Dicho en diferentes palabras: quien conoce se pone dentro de aquello que conoce, en este caso, dentro de un ambiente social.

Por su parte, Denise Jodelet (1984), seguidora del pensamiento de Moscovici, sostiene que las representaciones sociales son: “imágenes que condensan un conjunto de significados; sistemas de referencia que nos permiten interpretar lo que nos sucede e incluso dar un sentido a lo inesperado; categorías que sirven para clasificar las circunstancias, los fenómenos y los individuos con quienes tenemos algo que ver” (p. 472).

En síntesis, las representaciones sociales constituyen sistemas cognitivos en los que es posible reconocer la presencia de estereotipos, opiniones, creencias, valores y normas que suelen tener una orientación actitudinal positiva o negativa. Se componen, “a su

vez, como sistemas de códigos, valores, lógicas clasificatorias, principios interpretativos y orientadores de las prácticas, que definen la llamada conciencia colectiva, la cual se rige con fuerza normativa en tanto instituye los límites y las posibilidades de la forma en que las mujeres y los hombres actúan en el mundo” (Hoyos, 2013, p. 165).

3. Método

La investigación se desarrolló con un enfoque etnográfico desde un ángulo descriptivo e interpretativo, tomando como punto clave los significados construidos por la comunidad educativa para analizar las representaciones sociales en torno a la evaluación en el programa de Psicología de la Universidad Cooperativa de Colombia, sede Bogotá.

La etnografía es el estudio descriptivo de la cultura de una comunidad; el término es una denominación acuñada en el ámbito de los estudios antropológicos culturales. Para Anadón, “la etnografía es percibida como un enfoque que requiere inmersión directa del investigador en el medio estudiado con el fin de aprehender el “estilo de vida” de un grupo a partir de la descripción y la reconstrucción analítica e interpretativa de la cultura, de las formas de vida y de la estructura social del grupo estudiado” (citado por Bautista, 2011, p. 85).

La investigación etnográfica es un proceso de aproximación a una situación social, considerada de manera general en su propio contexto natural, método que permite lograr el objetivo fundamental de conocer las representaciones sociales que se han construido en torno a la evaluación de los aprendizajes en el contexto universitario, sobrepasando el discurso académico y accediendo a la comprensión empática del fenómeno objeto de estudio. Con este método, se pueden analizar y describir culturas y comunidades para acceder a las creencias y prácticas del grupo estudiado con el objeto de descubrir patrones y regularidades que surgen de la complejidad.

El referente de análisis fueron las representaciones sociales, utilizando técnicas como la *asociación libre*, para conseguir los significados espontáneos que emiten los estudiantes ante el término evaluación, y la *observación*, con el fin de examinar posturas, comentarios y actitudes ante el evento de diversos ejercicios evaluativos. Por último, el instrumento que más datos produjo fue la entrevista abierta que se realizó a estudiantes y docentes.

Al realizarse la investigación desde el programa de psicología, la población fue conformada por 60 estudiantes, 3 de cada semestre del diurno y 3 de cada semestre del nocturno, y 4 docentes del programa de psicología. Se incluyeron estudiantes de todos los semestres del diurno y del nocturno para conocer los criterios tanto de las personas que están iniciando la carrera como de los que ya han tenido algún proceso dentro de la misma. En cuanto a los docentes, se presentó la dificultad de contar con muy poca colaboración para la investigación.

4. Resultados y discusión

A continuación se exponen las concepciones de evaluación desde el punto de vista de los estudiantes y de los docentes donde se evidenciaron diferentes representaciones sociales derivadas del ejercicio evaluativo. Para llegar al reconocimiento de estas

representaciones, se plantearon unas categorías de análisis respecto a la evaluación, a saber: concepto, método, función y ética en el ejercicio de la evaluación.

4.1. Concepto

Los conceptos que tienen docentes y estudiantes de la evaluación se expresan bien en representaciones sociales que las reducen a un ejercicio memorístico, o bien como un modo de aprendizaje.

4.1.1. Evaluación como ejercicio de memorización

La práctica evaluativa no es vista por la mayoría de los estudiantes como una forma de medir para replantear y mejorar el aprendizaje, sino como una actividad donde lo que se considera es la capacidad de memorización: “pienso que la evaluación como la tenemos en este momento, mide la cantidad de conceptos que se quedan en nuestra memoria por un determinado tiempo (...), la mayoría de la gente estudia para los exámenes (...) y después lo olvidan (...)”. Este pensamiento demuestra que la evaluación no se considera un medio de valoración de una situación educativa, ni un proceso que permite llegar a una decisión favorable para el alumno.

Así, los estudiantes se sienten abordados con intención punitiva sin encontrar realmente la fortaleza de la evaluación como proceso de aprendizaje, que necesariamente incluye el enriquecimiento de la memoria, no en el sentido de memorización, sino como condición necesaria para que tenga lugar el aprendizaje; de hecho, aprender es guardar algo en la memoria. Las evaluaciones apuntan directamente a la capacidad de memoria; sin embargo, el concepto popular de la misma se manifiesta en la construcción de un rótulo negativo hacia la memoria reduciéndola a “recitar” puntualmente nombres, fechas y conceptos puntuales. No obstante, la memoria es la que va acumulando todo tipo de información nueva que conlleva la creación de competencias que requieren el uso de lo aprendido.

4.1.2. Evaluación como aprendizaje

El aprendizaje como resultado de la evaluación no es percibido por todos los estudiantes, ya que esta concientización se da a medida que la formación profesional va avanzando, como lo expresa una docente:

En los primeros semestres los estudiantes tienden a trabajar por la nota y no por el aprendizaje; sin embargo, en (...) 5º semestre en adelante, hay mayor tendencia a trabajar por el aprendizaje. (...) han alcanzado una conciencia mayor sobre la profesión o porque la asignatura les permite acercarse a una práctica educativa más aplicada. Por ende, la evaluación en la vida académica de los estudiantes tiende a tener mayor impacto en cuanto reconocen sus posibilidades, fortalezas y debilidades (...), además porque empiezan a actuar en equipos reales de trabajo y esto apoya y enriquece su proceso.

La docente reconoce que el desarrollo de la asignatura tiene una incidencia importante en la construcción de una representación social favorable de la evaluación, asumiendo la responsabilidad del profesor en la concientización que tiene el alumno frente a este ejercicio.

En los semestres más avanzados, los estudiantes tienen más conciencia de la fortaleza del ejercicio como señaló una alumna de 10º: “gracias a esos procesos de evaluación que continuamente se realizan, las cosas mejoran, es decir, se evalúa para mirar algo, intervenir en algo y se busca con ello generar una solución”. Los sujetos van llegando al verdadero concepto de la evaluación entendiéndola como una oportunidad para poner en práctica los conocimientos y como un proceso dinámico, abierto y útil para el

evaluado y el evaluador, conceptos que se acercan a los expuestos en las perspectivas teóricas de la evaluación de aprendizajes. Así, la evaluación se asume como una forma de aprendizaje, ya que el estudiante es consciente de que al ejercer su profesión será evaluado permanentemente por el desempeño que tenga como psicólogo, tanto por las instituciones como por los mismos clientes y pacientes.

4.2. Método en la evaluación

Al indagar sobre los métodos utilizados en la evaluación de los aprendizajes se encuentra una representación social de variedad de técnicas e instrumentos utilizados en el ejercicio.

4.2.1. Evaluación como variedad de instrumentos

En el proceso evaluativo la mayoría de los estudiantes reconocen diferentes instrumentos que se utilizan en la cotidianidad del aprendizaje, tales como exámenes, parciales, talleres grupales, control de lectura, exposiciones; otros entienden por evaluación únicamente el examen parcial o final; y unos pocos asumen que evaluación es la que ellos hacen a los docentes.

Los docentes entrevistados manifiestan que en su práctica utilizan diferentes instrumentos de evaluación pero observan que existe falencia en la comprensión lectora de los estudiantes o no leen los enunciados, por lo que en muchas ocasiones no siguen las instrucciones para el desarrollo de los ejercicios. Un docente señala que “un estudiante pierde una evaluación por diversas razones como la falta de comprensión lectora y análisis, no seguir instrucciones y el bajo compromiso con la asignatura”.

Los instrumentos utilizados en el programa de psicología están conforme a lo propuesto por Ronald Feo (2010), que resalta la conveniencia de utilizar diversos formatos que permitan recabar información apropiada a la temática en particular. En este sentido, se encuentran instrumentos de reflexión, memorísticos, investigativos, analíticos, descriptivos, de comprensión, de redacción, de observación, exposiciones, entre otros, los cuales son utilizados en su totalidad en la práctica evaluativa del programa de psicología de la Universidad Cooperativa.

Se reconocen, de forma general, dos grandes categorías de instrumentos, unos referentes al trabajo grupal y otros al trabajo individual. El trabajo grupal se evidencia en el desarrollo de talleres como respuesta al debate dado entre los integrantes del mismo; también predomina la oralidad en exposiciones, mesas redondas y análisis de casos. La prueba oral es privilegiada por muchos estudiantes porque pueden dar explicación desde su punto vista y ser retroalimentados simultáneamente, desplegándose una modalidad de coevaluación automática que fomenta el pensamiento crítico.

En la evaluación individual predomina la expresión escrita y los instrumentos memorísticos, impulsando el sentimiento de vigilancia y las emociones de ansiedad e inseguridad. El método escrito individual tiene un enorme impacto en los estudiantes llegándose a reducir el concepto de evaluación al instrumento de tipo memorístico y punitivo donde el estudiante se percibe como un sujeto pasivo que responde a un evaluador quien ha definido los parámetros de la prueba, el tiempo y la forma de calificar. Es por esta práctica que se ha construido una representación social de la evaluación como formulario sin posibilidad de mostrar pensamiento creativo y crítico.

4.3. Función de la evaluación

En la categoría de función de la evaluación se encontraron tres representaciones sociales: retroalimentación; calificación y requisito; e imposición.

4.3.1. Evaluación y retroalimentación

Los estudiantes exponen que algunos docentes hacen retroalimentación de los ejercicios, lo que convierte la evaluación en un método de aprendizaje y un modo en que el alumno puede observar las fortalezas y debilidades de su formación profesional: “me permite ver en lo que estoy fallando y en lo que tengo fortalezas”; así, un estudiante de 9° semestre considera esta práctica importante para medir el conocimiento que ha adquirido en determinado tema, reconociendo la función diagnóstica de la evaluación.

No obstante, para otros la evaluación es una actividad normativa que los profesores aplican únicamente por cumplir un requisito y no explican la razón de la calificación registrada: “lo hacen como requisito para calificar”. Un estudiante que piense de esta manera también puede considerar que realiza sus evaluaciones para cumplir el requisito. Así, se desconocen las funciones expuestas por Gimeno Sacristán (1996) referentes al rol pedagógico de la evaluación, a la proyección psicológica, al poder de control y a la organización educativa.

Como afirman Osorio y López Mendoza (2014), la evaluación es un proceso que supervisa la instrucción y la misma no se ubica en ningún momento instruccional en particular, pues se entiende que el monitoreo y la retroalimentación con fines instruccionales son constantes. Es así que la retroalimentación debe formar parte del interés formativo tanto del docente como del estudiante.

Algunos docentes afirman que, cuando el alumno recibe de vuelta un trabajo, lo que le interesa es la calificación obtenida y, en ese sentido, los comentarios que registra el profesor pierden interés. Así, el *feedback* tiene poca eficacia si el énfasis está puesto en la calificación, lo que desanima al docente, que no observa las bondades del ejercicio a la vez que el estudiante presume que avanzar implica dejar atrás lo que ya ha sido calificado.

4.3.2. Evaluación como calificación y requisito

Para muchos estudiantes la evaluación cumple el objetivo de otorgar una nota y se aleja de la actividad del aprendizaje, por lo que la calificación se torna el objetivo principal del ejercicio en el aula, ya que es la que permite que la persona sea promocionada a un curso siguiente. Así se entiende en estas palabras de un estudiante de 1^{er} semestre: “uno como estudiante puede utilizar unos métodos o lo que sea para alcanzar una nota para cumplir un requisito para ser evaluado sin necesidad de tener aprendizaje”. Visto de este modo, la evaluación es una práctica vacía cuyo único objetivo es proferir calificaciones.

Como consecuencia de entender la evaluación como requisito, se piensa que este ejercicio se debe cambiar y no aplicarlo solo por norma y como medio de calificación, como lo manifiesta un estudiante de 5° semestre: “yo pienso que, para mejorarlo, se debería evaluar en términos de procesos y en una dinámica discursiva, no limitarse a colocar una nota sino convertirla en un discurso, en donde el estudiante intervenga”. Este es uno de los casos en que un estudiante demuestra la necesidad de cambiar las

prácticas evaluativas para que no se limite a la obtención de un valor numérico que refleje lo aprendido, sino que represente los elementos que hay que reforzar en el aprendizaje del tema. El sujeto se acerca mediante su experiencia a entender la evaluación como un proceso de diálogo, comprensión y mejora, como discusión y reflexión compartida de todos los que están implicados en la actividad evaluada, conforme lo expone en sus investigaciones Santos Guerra (1993).

4.3.3. Evaluación como imposición

Los enfoques teóricos sobre evaluación de los aprendizajes van dirigidos al docente e implícitamente se le otorga un papel pasivo al estudiante, lo que ha creado una representación social de imposición o “molestia necesaria” para el alumnado. La pasividad del estudiante ante la formación que se desprende de la evaluación se observa de manera reiterada en el ámbito universitario. Así como el docente no siempre utiliza la evaluación como mecanismo de enseñanza, el estudiante tampoco se libera de una actitud reduccionista frente a este ejercicio y no vuelve sobre los resultados para mejorar su aprendizaje: “el estudiante se queda con lo que dice el docente, pero no se dirige a los libros, no practica en otros lados”. Este comentario en 8º semestre demuestra que algunas personas reconocen la pasividad de quienes están en formación al depender únicamente de los datos que reciben en las aulas, sin buscar ninguna forma de ampliar ese conocimiento.

La pasividad contradice un principio de la formación profesional y al tenor de la misión institucional de la Universidad Cooperativa referente a la reflexión teórica y práctica que permite acceder a la autonomía en el aprendizaje. Se predica la orientación hacia el pensamiento crítico y la búsqueda de transformaciones en los ámbitos social, cultural, educativo, tecnológico, científico y familiar. La evaluación en la educación superior se debe constituir en un eje de análisis del modelo positivista imperante, ya que pone en juego tanto las potencialidades del estudiante universitario como sus limitaciones, todo lo cual es una motivación para el razonamiento en aras a movilizar discursos y paradigmas estáticos y hegemónicos que coartan el dominio del propio comportamiento, de los pensamientos y de las emociones.

4.4. Ética

La categoría ética surgió especialmente de la observación y la técnica de asociación libre, lo que resaltó algunas características sociales que se han implantado como representaciones de modalidades para interactuar en el mundo. Aquí se encontró el imaginario social sobre el uso del poder, el plagio y la violencia como recurso.

4.4.1. La evaluación como ejercicio de poder

En el tema de la evaluación surge la representación social respecto al poder del docente tanto en el buen entendimiento del término, como en la utilización corrupta del mismo. Así mismo, el estudiante evidencia su propio manejo de poder cuando traslada su responsabilidad a la institución o al docente, cambia de opinión y de comportamiento según su conveniencia, invoca razones lógicas pero descontextualizadas para encubrir sus demandas, exige la perfección de los demás y exagera los errores, no admite sus propios incumplimientos en la presentación de trabajos, se victimiza por su incapacidad de interesarse por el aprendizaje o amenaza de forma encubierta para que el docente acceda a subir la nota.

La victimización estuvo presente en la historia colombiana desde la Conquista y la Colonia, donde se violentaron los derechos de los nativos. Esta situación ha pasado a la

memoria colectiva como un recurso estratégico que demuestra un manejo de poder del individuo donde se asume como sujeto pasivo, atemorizado, dependiente y sumiso ante un supuesto dominio total del otro. En muchas situaciones cotidianas, el victimista se disfraza de víctima, consciente o inconscientemente, simulando una agresión o menoscabo inexistente y responsabilizando erróneamente al entorno o a los demás.

La autovictimización es una actitud que consiste en considerarse como el afectado por ciertas situaciones específicas y siempre ser el doliente de las circunstancias. En esta postura el individuo es incapaz de reconocer el mérito ajeno y considera que las demás personas deben sentir una pena profunda y un dolor solidario, que si no es así demuestran poca humanidad. Es por eso que esta estrategia se utiliza para probar la bondad y empatía del docente mediante un ablandamiento en la exigencia de la prueba o el rompimiento de normas que rigen el trabajo evaluado y que fuera acordado previamente.

Es evidente que es más fácil ponerse en la posición de víctima cuando el individuo es culpable de su incumplimiento, demostrando incapacidad para medir sus acciones y calcular las consecuencias de las mismas, y prefiere buscar la solución del problema adjudicándoles la responsabilidad a otras personas. Cuando el estudiante le muestra a los distintos docentes lo mal que está, se imagina que esta actitud se impregna en ellos convirtiéndose en una estrategia apropiada para enfrentar fallas demostradas en el acto evaluativo.

Por su parte, el docente también ejerce su poder en el tema aquí tratado; ejerce su dominio sobre lo que evalúa y cómo lo evalúa, incluso sobre la posibilidad de convertir esa evaluación en ejercicio de aprendizaje, de ofrecer la oportunidad o no de mejorar la calificación del alumno. No obstante, en su discurso tiene la percepción de perder cada vez más el dominio sobre su actividad, sintiéndose acosado, descalificado y culpado por alumnos, por los mensajes de la institución y por la sociedad en general, respecto al comportamiento de la comunidad estudiantil como resultado de una supuesta labor deficiente. No obstante, existe una representación del poder del docente en la población de estudiantes, de los mismos docentes y de la sociedad en general en el manejo de la evaluación de los aprendizajes.

El manejo abusivo del poder por parte del docente y del estudiante redundan en indiferencia. El desánimo del docente unido a la apatía del estudiante convierte la práctica evaluativa en un proceso aséptico, desprovisto de poder fructífero y de contenido ideológico. La discusión sobre el tema es pobre aunque conlleva al discurso ético de una educación sin voluntad que es la que finalmente determina las técnicas y los procedimientos y el sentido que se le otorga a los mismos. Quienes ejercen el poder desde el deseo y no desde el beneficio del otro generan tensiones y conflictos que deconstruyen el componente formativo de la evaluación.

4.4.2. El plagio en el ejercicio evaluativo

En la cultura estudiantil es común recurrir al plagio por medio del cual el estudiante pretende obtener calificaciones aprobatorias sin esforzarse; de esta manera la evaluación se convierte en un elemento de vicio y corrupción ratificándose la supremacía del valor de la nota sobre el aprendizaje.

Un reconocido político dijo alguna vez “tenemos que reducir la corrupción a sus justas proporciones” (Rojas, 2011). Con estas palabras se reconoce que la corrupción es una expresión de los vicios culturales que se han aprendido en un país donde la violación de

los derechos ha sido una constante a través de la historia, expresada en las acciones de las personas en diferentes espacios, incluyendo el medio académico donde se evidencia a través del plagio como una perpetuación de la *cultura del vivo*.

La facilidad para acceder a cualquier información globalizada se ha convertido en una tentación constante para las personas que encuentran allí todo tipo de datos y reflexiones ya elaboradas que son compatibles con el lema estudiantil del “menor esfuerzo”, afianzando una tendencia a la corrupción en el evento de tener que realizar ejercicios tediosos que otrora fueran parte de la rutina estudiantil, pero en la actualidad son motivo de descontento porque el tiempo de las personas se ha copado con las ocupaciones que ha traído la implementación de las TIC en la vida cotidiana y en la interacción social.

Con estas acciones se muestra que en la práctica la evaluación se puede deformar hasta niveles no previstos en teorías y normas, lo que hace indispensable una revisión de las conductas que suceden alrededor de los diferentes actos humanos que nunca están totalmente ligados a lo estipulado por instituciones, leyes y preceptos sociales. Esto lleva el tema de la evaluación a ser analizado desde los principios de los derechos humanos y la seguridad.

4.4.3. *La violencia como recurso*

Dentro del estudio adelantado se encontró a dos estudiantes a quienes denominaremos HO quienes a lo largo de su carrera cometieron toda clase de infracciones y acciones temerarias contra docentes, compañeros de estudio, directivos y personal administrativo, lo que puso en evidencia la poca preparación que tiene la institución ante casos extremos de anomia y anarquía que llegan a alcanzar los límites de la delincuencia.

Los estudiantes HO luego de su primera clase de la carrera iniciaron un liderazgo negativo y desde ese momento mantuvieron una actitud de protesta por todo lo que estaba a su alrededor. Los docentes y directivos respondían solícitamente a las peticiones, pero en poco tiempo esa actitud de desaprobación se transformó en vanagloria con la consecuente exigencia temeraria e irrespetuosa a todo docente que les reprobara sus incumplimientos en asistencia, tareas y ejercicios evaluativos.

Con el protagonismo desarrollado por HO, se fue mostrando ante el grupo de estudiantes el poder que tiene la agresividad y se fueron generalizando actitudes irreverentes entre un grupo de alumnos que no admitían la pérdida de ningún ejercicio evaluativo. Así se vio dentro del claustro universitario cómo se construía una representación social de la violencia como medio de solucionar las dificultades. La conducta belicosa elimina la función formativa de la evaluación colocando como fin último un certificado sin sustento de competencias. Lo que se juega en la anulación de las bondades del ejercicio evaluativo son los principios básicos de humanidad, de respeto y de reconocimiento a la labor docente, de la imagen institucional y de la ética profesional.

Los estudiantes HO durante su proceso protagonizaron toda clase de acciones violentas para conseguir las calificaciones aprobatorias en todas las asignaturas. Para completar el historial de conducta abusiva, se anexaron a un trabajo de grado que estaba adelantando otro estudiante y, al tener total incapacidad de relación pacífica, terminaron en conflicto con el estudiante, el asesor y toda persona que tratara de aplicar las normas del reglamento. Finalmente, estos estudiantes se graduaron con el

trabajo del compañero, quien dejó de asistir a la universidad por temor a ser atacado. Así fue que esta graduación se logró mediante múltiples conflictos y agresiones no solo de HO sino de sus familiares y amigos que llegaban al campus con amenazas y escándalos para que se les suspendieran los procesos disciplinarios que tenían en curso por diversos motivos.

En este caso se demostró que en algunos estudiantes predomina la creencia de que la violencia es un recurso que conlleva al fortalecimiento de su identidad social, por lo que exhiben actitudes desafiantes como la agresión y la violación de las normas, lo cual no puede verse como un caso aislado, pues al examinarlo desde el prisma de las representaciones sociales se observa que la situación ha visibilizado pensamientos propios del sentido común de una sociedad que ha integrado en sus saberes culturales la imposición y la temeridad como un camino para asegurar sus derechos, olvidando los límites que afectan en una escala mayor los derechos humanos del entorno social.

Los medios violentos de expresión fueron rechazados por la mayoría de los estudiantes, quienes expresaron su descontento en relación a que a los agresores se les admitían todas las faltas mientras que a quien se acogía a la norma se le exigía mayor esfuerzo. Los argumentos que expone la comunidad estudiantil no son otros que ratificar la representación social respecto al predominio de la fuerza, ya que nunca se propuso un apoyo por parte de la comunidad para que prevaleciera el respeto a los derechos humanos, el conducto regular y el seguimiento de la norma como medio de recuperar la estabilidad institucional.

Los estudiantes violentos saltan para sí mismos el derecho fundamental del “debido proceso”, puesto que buscan obtener la calificación aprobatoria de forma inmediata y con argumentos no válidos que solo fortalecen la temeridad y la osadía. El marco legal coloca el tema de la evaluación del aprendizaje y el desempeño humano en un rol determinado por el contexto de la ética de mínimos, cuyo principal referente es la Declaración Universal de los Derechos Humanos divulgados por la ONU en 1948. Desde el mismo principio filosófico, ético y axiológico, la Constitución Política de Colombia de 1991 en su artículo 67 define la concepción del Estado sobre la educación así: “La educación es un derecho de la persona y un servicio público que tiene una función social (...). La educación formará al colombiano en el respeto a los derechos humanos, a la paz y a la democracia”. Este principio constitucional se viola en el evento de asumir actitudes violentas contra cualquier sujeto de la comunidad educativa.

El problema radica en no asumir una actitud autoevaluativa que permita entender la necesidad de fortalecer la estructura profesional, actitud que no cambia al momento de desarrollar su trabajo de grado, buscando plagiar otros textos o acudir a los “carteles de tesis” donde venden el informe de investigación que solicita el estudiante (Benavides, 2012). Predominan así unos imaginarios que cada vez toman más fuerza en la población universitaria como el “todo vale” o “la ley es para violarla”.

El trabajo de grado es la evaluación más importante que permite validar las competencias profesionales del estudiante, pero al ser todo sujeto de negociación ha dado lugar a la conformación de una zona negra en el ámbito educativo que abre una brecha entre la evaluación formativa, la formación ciudadana y la aptitud profesional.

5. Conclusiones

Es asumiendo y no negando la complejidad del mundo de la vida que se hace necesario hacer una permanente reflexión sobre el factor ético de la evaluación del aprendizaje para que no se desdibuje el proceso formativo bajo acciones que se van implementando como forma rápida de solución de problemas que necesariamente están siempre presentes en toda actividad humana.

El manejo inadecuado del poder del docente y de los estudiantes en el ejercicio evaluativo conlleva al fortalecimiento de la injusticia que se enraiza en unas prácticas culturales, de ahí que es indispensable hacer un llamado a la comunidad universitaria para que toda actividad evaluativa se sustente los principios de Justicia social expuestos por Murillo e Hidalgo (2015): poseer una cultura de lucha por la Justicia Social; fomentar del desarrollo integral de los estudiantes basado en el respeto; desarrollar procesos académicos justos; reconocer la educación como un proceso vinculado con la raza, el género, la cultura, la orientación sexual y la capacidad del estudiante; tener altas expectativas hacia los estudiantes; contar con un liderazgo democrático y distribuido; fomentar la colaboración con el entorno social; y favorecer el pensamiento crítico y democrático.

La evaluación siempre encierra un juicio de valor y debe atender a diversos criterios en la formación del estudiante, quien a partir de la misma observará el desarrollo de sus capacidades en el ser, en el saber, en el hacer y en el convivir, por lo que exige no solo una adecuada planeación sino también una retroalimentación y la oportunidad de aprender y corregir desde el concepto recibido.

En el proceso evaluativo confluyen pensamientos, sentimientos, impulsos, costumbres y simbologías que convierten el espacio regulado por normas y perspectivas teóricas en una dimensión confusa y llena de prácticas que se alejan del discurso oficial. La evaluación en la educación superior es un tema en el que interviene la cultura y no solo debe ser vista desde el juzgamiento al docente, como se ha concebido tradicionalmente y lo manifiesta a manera de crítica en una entrevista Manuel Gil: “el prejuicio radica en considerar que, si no la única, por lo menos sí la principalísima causa de los problemas de los problemas del aprendizaje en el país es consecuencia de la incapacidad, la flojera, e incluso el fenotipo del personal docente” (Ordorika, Gil, Rodríguez y Rueda, 2016). Esto lleva a perpetuar la ausencia del estudiante en la optimización de la actividad evaluativa y buscar soluciones limitadas y sin trascendencia.

Analizar la postura de los actores sociales más incidentes en el proceso evaluativo nos lleva a una instancia más amplia que es la institución universitaria, a la que le corresponde traspasar el texto normativo para crear mecanismos que permitan mantener en el lugar que corresponda las acciones de docentes y estudiantes, especialmente cuando se observa la desviación de la norma.

La universidad tiene un compromiso frente a la sociedad de formar profesionales competentes para responder a las dinámicas del mundo, no solo contribuyendo a la construcción y difusión del conocimiento y apoyando el desarrollo competitivo del país, sino también buscando el mejoramiento de la calidad de vida de las comunidades (Universidad Cooperativa de Colombia). Por ello, en concordancia con lo expuesto por el docente investigador Ricardo Gaete-Quezada (2015), es “fundamental atender adecuadamente las demandas internas mediante el correcto uso del poder y autoridad,

resolviendo conflictos y generando acuerdos internos necesarios para mantener la estabilidad institucional”. (p. 8).

Es importante que la Universidad se proponga transformar las representaciones sociales imperantes sobre la evaluación como requisito o como mecanismo de poder para que se la conciba siempre como autoevaluación, entendiendo que el concepto del sí mismo se fundamenta en los mensajes recibidos del otro. Como lo afirma George Mead (1972), la sociedad es interacción y “funciona como un equipo y no es la acción individual aislada la que se impone, sino la comunicación y permanente retroalimentación que el individuo recibe del grupo social” (p. 15), lo que se convierte en autoconcepto.

Referencias

- Álvarez, J. (2001). *Evaluar para aprender, examinar para excluir*. Madrid: Morata.
- Araya, S. (2002). *Las representaciones sociales. Ejes teóricos para la discusión*. San José de Costa Rica: FLACSO.
- Bautista, N. (2011). *Proceso de la investigación cualitativa*. Bogotá: Manual Moderno.
- Bautista, N. (2015, 27 de abril). *Investigación cualitativa: Análisis de caso*. Recuperado de <https://prezi.com/vl1g2-vr0an1/investigacion-cualitativa-analisis-de-caso/>
- Benavides, J. (2012, 2 de noviembre). *¿Cuánto cuesta una tesis?* Recuperado de http://www.cartelurbano.com/content/precio-de-una-tesis#.U_9B5_15Mn4
- Correa, S., Puerta, A. y Restrepo, B. (2002). *Investigación evaluativa*. Bogotá: ICFES.
- Feo, R. (2010). Orientaciones básicas para el diseño de estrategias didácticas. *Tendencias Pedagógicas*, 16, 220-236.
- Gaete-Quezada, R. (2015). El techo de cristal en las universidades estatales chilenas. Un análisis exploratorio. *Revista Iberoamericana de Educación Superior*, 6(17), 3-20.
- Gimeno Sacristán, J. (1996). *Comprender y transformar la enseñanza*. Madrid: Morata.
- Hoyos, C. (2013). Representaciones sociales en el adolescente sobre la norma y el delito. *Advocatus*, 21(21), 161-172.
- Jodelet, D. (1984). La representación social: Fenómenos, concepto y teoría. En S. Moscovice (Comp.), *Psicología social II: Pensamiento y vida social* (pp. 469-493). Barcelona: Paidós.
- Mead, G. (1972). *Espíritu, persona y sociedad*. Buenos Aires: Paidós.
- Mora, M. (2002). La teoría de las representaciones sociales de Serge Moscovici. *Athenea Digital*, 2, 1-25.
- Moreno Olivos, T. (2009). La evaluación del aprendizaje en la universidad: Tensiones, contradicciones y desafíos. *Revista Mexicana de Investigación Educativa*, 14(41), 563-591.
- Moscovici, S. (1979). *El psicoanálisis, su imagen y su público*. Buenos Aires: Huemul.
- Murillo, F. J. e Hidalgo, N. (2015). Enfoques fundamentantes de la Evaluación de Estudiantes para la Justicia Social. *Revista Iberoamericana de Evaluación Educativa*, 8(1), 43-61.
- Orden, A. de la (1981). Evaluación del aprendizaje y la calidad de la educación. En V. García Hoz (Comp.), *La calidad de la educación. Exigencias científicas y condicionamientos individuales y sociales* (pp. 111-132). Madrid: CSIC.

- Organización de Naciones Unidas. (1948). *Declaración Universal de Derechos Humanos*.
- Osorio, K. L. y López Mendoza, A. (2014). La retroalimentación formativa en el proceso de enseñanza-aprendizaje de estudiantes en edad preescolar. *Revista Iberoamericana de Evaluación Educativa*, 7(1), 13-30.
- Proppe, O. (1990). La investigación de la evaluación como una forma de potenciar el desarrollo en las escuelas y el profesionalismo en los profesores. *Revista de Educación*, 293, 325-345.
- Rojas, J. (2011). Izquierda en Colombia: ¿Destinada a la muerte política? *El Tiempo*, 1-3.
- Ordorika, I., Gil, M., Rodríguez, R. y Rueda, M. (2016). Reforma educativa y evaluación docente: El debate. *Perfiles Educativos*, 38(151), 190-207.
- Ruiz, Y. A. (2014). Evaluación del aprendizaje: Aproximación conceptual. *Aula Magna*, 2, 1-14.
- Santos Guerra, M. Á. (1993). La evaluación: Un proceso de diálogo, comprensión y mejora. *Investigación en la Escuela*, 20, 23-35.

Funcionamiento Diferencial del Ítem en una Evaluación Estandarizada según Necesidades Educativas Especiales Transitorias

Differential Item Functioning for Students with Special Educational Needs in a Primary Education Standardized Test

Fabiola Silva Morales y María Verónica Santelices Etchegaray*

Pontificia Universidad Católica de Chile

Un aspecto fundamental en la evaluación de aprendizajes es la validez de los resultados de la prueba utilizada, que no solo supone medir lo que se dice sino considerar las implicaciones de esos resultados, porque a partir de ellos se toman variadas decisiones. La existencia de funcionamiento diferencial del ítem (DIF) indica que sujetos que tienen igual grado de habilidad presentan diferente probabilidad de responder correctamente a un ítem, lo que se relaciona con el sesgo, es decir, favorecer a un grupo sobre otro en la evaluación. En los últimos años, la integración y el aprendizaje de los alumnos con necesidades educativas especiales transitorias (NEET) han sido el foco de políticas educacionales especiales que involucran recursos, especialistas y nuevas formas de trabajar en las escuelas de Chile. Mediante un enfoque metodológico secuencial cuantitativo se exploró el funcionamiento diferencial del ítem de una prueba estandarizada aplicada por un centro de medición en educación (MIDE UC) a más de 100 establecimientos en Chile en las áreas de lenguaje y matemática en los niveles desde 4º a 8º Básico para niños con y sin NEET. Un primer análisis cuantitativo exploró el funcionamiento diferencial del ítem (DIF) obteniéndose similitudes en la habilidad promedio de los grupos y pocos ítems con DIF significativo (menos del 10%). Posteriormente, se exploró la forma en que los colegios operacionalizan la variable integración, observándose que el grupo de niños sin NEET usado en el análisis estadístico puede haber sido, en realidad, un grupo híbrido que incluía a algunos niños con NEET.

Palabras clave: Evaluación, Sesgo, Funcionamiento diferencial del ítem (DIF), Sesgo estadístico, Integración educativa, Validez, Necesidades educativas especiales transitorias (NEET).

An important aspect of learning assessment is the validity of the results of the test used because they are the basis for social, political or institutional decisions. The existence of differential item functioning (DIF) indicates that subjects having presented the same degree of skill, have different probability of correctly respond to an item, which strongly affect an aspect of the validity of the test, which relates to the bias, i.e. favor one group over another in the evaluation. In recent years, integration and learning of students with temporary special educational needs (NEET) have been the focus of special education policies involving resources, specialists and new ways of working in schools in Chile. Using a sequential quantitative methodological approach we explored the differential item functioning, aspect of the validity of the results, of a standardized test administered by measurement center (MIDE UC) to more than 100 schools in Chile in language and mathematics from 4th to 8th grade to children with and without NEET. An initial analysis of differential item functioning (DIF) showed similarities in the average group ability and few items with significant DIF (less than 10%). Later, we explored how schools operationalize the variable depicting NEET. Results suggest that the group of children without NEET used in the statistical analysis may have been, in fact, a hybrid group including some children with NEET.

Keywords: Assessment, Bias, Differential item functioning (DIF), Statistical bias, Educational integration, Validity, Temporary special educational needs.

*Contacto: vsanteli@uc.cl

1. Introducción¹

La evaluación del aprendizaje es una herramienta de recolección de información respecto de los saberes de los alumnos tanto a nivel de aula como a gran escala. Dicha información determinará decisiones que involucran a variados agentes (Navas, 2012). La medición se debe realizar a través de instrumentos que aseguren criterios de calidad en las interpretaciones de los resultados obtenidos (Prieto y Delgado, 2003). La validez ya no es solo medir lo que se pretende, sino que, según Prieto y Delgado (2003), corresponde al grado en que las interpretaciones y los usos que se hacen de las puntuaciones están justificados científicamente, coincidiendo con los estándares internacionales para las pruebas educativas y psicológicas establecidos por la American Educational Research Association (AERA), la American Psychological Association (APA) y el National Council on Measurement in Education (NCME). Es así como el proceso de validación de una prueba involucra la acumulación de evidencia para la interpretación y la relevancia de las puntuaciones de la prueba. Un aspecto de la validez de la prueba aplicada implica la exploración de la presencia de funcionamiento diferencial del ítem (DIF por su sigla en inglés) donde se distinguen dos grupos que son comparados, uno llamado de referencia (generalmente mayoritario) y el otro, focal, donde se sospecha que las propiedades psicométricas de los ítems puedan tener valores distintos (Gómez-Benito, Hidalgo y Guilera, 2010). La existencia de DIF es un antecedente relevante e indicativo de la posible existencia de sesgo. El sesgo es un “error sistemático originado por deficiencias en el test o en el modo en que éste es usado, que produce una distorsión en el significado de las puntuaciones y que adultera la interpretación propuesta” (Elosúa, 2003, p. 316). El sesgo representa un problema de equidad, ya que la puntuación obtenida por un sujeto en un test solo debe depender de su aprendizaje y no de sus características personales como etnia, género o necesidades educativas.

Desde 2010, Chile ha implementado políticas que promueven la inclusión de alumnos con necesidades educativas especiales permanentes (NEEP) y transitorias (NEET) proporcionando diversidad en las aulas. Sin embargo, la evaluación de aprendizajes que se realiza a nivel nacional no ha incorporado este aspecto en sus estándares de calidad.

El objetivo de esta investigación es analizar el DIF como un aspecto de la validez de una prueba estandarizada llamada prueba SEPA (Sistema de Evaluación para el Aprendizaje), aplicada en Chile por el centro de medición en evaluación de la Universidad Católica, MIDE UC, en las áreas de lenguaje y matemática desde 4° a 8° Básico, según la variable NEET. En ella, se aborda la interrogante respecto de si el puntaje obtenido en estas pruebas es reflejo del grado de habilidad que el alumno posee y si está libre de sesgo, es decir, es independiente de la existencia de necesidades educativas especiales transitorias (NEET).

Para complementar los resultados obtenidos, se exploró la forma en que los colegios que aplican la prueba SEPA reportan la variable NEET, indagando respecto de cómo se determina dicha clasificación, entre otros.

¹ Proyecto financiado en parte por el proyecto ANILLO SOC1107 de Conicyt Chile.

2. Fundamentación teórica

2.1. Inclusión

Durante las últimas décadas ha crecido la preocupación mundial por el derecho a la educación planteándose la necesidad de inclusión de personas con todo tipo de discapacidad en los sistemas educativos, generándose acuerdos internacionales en los que Chile participó (Organización de las Naciones Unidas, 2008). Es así como se genera un marco legal en torno a la definición y forma de trabajo de las NEET en los establecimientos del país. La inclusión escolar en Chile insta a cambiar el modo en que se concibe la discapacidad, desde una visión inicial centrada en la corrección de un defecto hacia una mirada donde la necesidad educativa es concebida como “resultado de la interacción entre los recursos, las carencias de los individuos y el medio” (Infante, Matus y Vizcarra, 2011) siendo parte del contexto existente, no arraigado en el déficit de la persona sino en las barreras sociales y culturales para evidenciar sus verdaderas capacidades (Tenorio, 2011).

De acuerdo con lo descrito, el diagnóstico de una NEET en establecimientos con financiamiento público en Chile es realizado por un equipo multidisciplinario de profesionales, que aplican una batería de pruebas estandarizadas y evalúan en varios ámbitos al estudiante, incluyendo lo cognitivo, afectivo-social, hábitos de estudio, familiar, salud y nutrición (Decreto 170/09). El marco legal permite que hasta el 20% de alumnos en aula presenten NEE, 7% permanentes y 13% transitorias. Si el número de alumnos por sala es en promedio 40, existirían 5 estudiantes con NEET y 2 con NEEP por aula. De no existir alumnos con NEEP, se permite la incorporación al proyecto de integración escolar (PIE) de un número máximo de 7 alumnos por aula.

El PIE es una estrategia inclusiva del sistema escolar chileno cuyo propósito es entregar apoyos adicionales a los estudiantes que presentan NEE permanentes o transitorias contribuyendo al mejoramiento de la calidad de la educación que se imparte en los establecimientos favoreciendo la presencia, la participación y el logro de los aprendizajes esperados (Decreto 170/09). A través del PIE, se entregan recursos a las escuelas para que los profesionales idóneos asistan pedagógicamente a los alumnos con necesidades educativas en el aula. Las escuelas, a su vez, se comprometen a promover una cultura que incentiva la inclusión y respeta la diversidad en el aula.

Según estadísticas del Ministerio de Educación (2013), un 6,3% de niños en Chile presentan NEE, concentrándose en el sector municipal y particular subvencionado. El 70% de ellos corresponde a NEET, las cuales se distribuyen en dificultades específicas de aprendizaje (DEA, 38%), trastorno específico del lenguaje (TEL, 21%), trastorno de déficit atencional (TDA, 18%) y coeficiente intelectual en rango limítrofe (FIL, 23%).

2.2. Adecuaciones

En las pruebas estandarizadas es posible que dadas las condiciones de aplicación del test se produzcan diferencias en los resultados que no sean propias de las diferencias en la habilidad. Para evitar esto, se realizan algunas adecuaciones que permitan a los estudiantes con NEET evidenciar su verdadero grado de aprendizaje (Sireci, Scarpati y Li, 2005). Varios investigadores han trabajado en esta línea, estudiando los efectos de la extensión del tiempo de aplicación del test, su presentación, el uso de ciertos materiales y equipos (calculadora) o la explicación en voz alta de algunos ítems y un ambiente de prueba tranquilo (Cahalan-Laitusis, 2010; Cawthon, 2010; Cohen, Gregg y Meng, 2005; Kettler et al., 2005; Mandinach, Bridgeman, Cahalan-Laitusis y Trapani, 2005).

En este sentido, Cohen et al. (2005) concluyen que las diferencias encontradas en el desempeño de alumnos con NEET y sin NEET no tienen su origen en la existencia de la necesidad educativa, sino que se explicarían por la falta de conocimiento del alumno, y que una extensión de tiempo solo beneficiaría al alumno con NEET que posee la habilidad matemática. Kettler et al. (2005) comparan el desempeño promedio entre niños sin NEET y con NEET encontrando que los alumnos con NEET presentan menores puntajes promedio sin adecuaciones y que, cuando se implementan, mejoran su desempeño.

Profundizando en el uso de adecuaciones, Mandinach et al. (2005) relacionan el impacto del tiempo extendido y la posibilidad de que la prueba se separe en secciones con el desempeño obtenido en el test SAT de lenguaje y matemática de alumnos con y sin NEET. Este estudio coincide con lo reportado por Cohen et al. (2005), planteando que las respuestas incorrectas se deben a la falta de conocimiento del alumno y no a la falta de tiempo. Sin embargo, contrasta con lo encontrado por Kettler et al. (2005), quienes verifican que la extensión de tiempo mejora los rendimientos en matemática más que en lenguaje, afectando principalmente a estudiantes con nivel de habilidad medio y/o alto, pero no habría ningún efecto cuando el estudiante posee bajo nivel de habilidad.

Mientras que algunos antecedentes muestran que los niños con NEET presentan menor rendimiento promedio en lenguaje y matemática (Kettler et al., 2005; Mandinach et al., 2005), Gregg (2007) comprueba que los estudiantes con NEET presentan rendimiento académico comparable a sus pares. En esta línea, Meltzer (2006) encuentra que los estudiantes con NEET solo requieren, en promedio, un semestre más que los alumnos sin NEET para cursar sus estudios superiores. Ello indicaría que no existen diferencias en la habilidad promedio entre los estudiantes con NEET y sin NEET.

Otros estudios exploran el menor rendimiento promedio de niños con NEET comprobando que estos presentan mayores índices de soledad y menores índices de autoeficacia académica (Idan y Margalit, 2014) y menor autorregulación en las tareas escolares y horas de estudio (Hen y Goroshit, 2014), lo que influiría en su rendimiento.

2.3. Estudios de DIF que involucran NEET

La mayoría de las investigaciones que relacionan DIF con NEET exploran el efecto de las adecuaciones sobre el desempeño demostrado en lenguaje y matemática. En ellas, el ajuste más frecuente y eficiente en lenguaje es extender el tiempo de aplicación. En matemática, en tanto, se encontró que algunos estudiantes con NEET mejoran su rendimiento cuando se realiza una explicación en voz alta de los ítems (Sireci et al., 2005).

Cho, Lee y Kingston en 2012 analizaron DIF según NEE en una prueba de matemática e historia aplicada a niños de 3° a 8° y encontraron un 21% de ítems con DIF con una tendencia mixta. Además, no hallaron diferencia en la habilidad promedio entre los grupos. Adicionalmente, el estudio muestra que algunos alumnos asignados con NEE respondían como alumnos sin NEE.

En 2013, Finch y Finch analizaron DIF en una prueba de lenguaje y matemática aplicada a alumnos de 3° de primaria en Estados Unidos, correlacionando multidimensiones (adecuación, escuela, género, etnia, raza, etc.) en un contexto de múltiples niveles. Los hallazgos reportados muestran que la existencia de DIF puede ser multicausal. Este argumento es compartido por el trabajo de Lin y Lin (2014)

quienes consideran que el DIF se explica mejor considerando la posible interacción entre variables tales como género, actitud hacia el aprendizaje, adecuación y habilidades.

En Latinoamérica existen pocos estudios que aborden DIF según NEET. Un aporte es la investigación de Moreira (2008), que comprueba la presencia de DIF en una prueba de matemática aplicada a alumnos de enseñanza media según trastorno de déficit atencional con hiperactividad (TDAH), encontrando que la prueba presenta un 18% de ítems con DIF de tamaño irrelevante pero estadísticamente significativo. Un análisis de los ítems con DIF indica que presentan vocabulario impreciso y redacción poco clara, entre otros, lo que perjudicaría a los niños con TDAH.

Dado que las políticas públicas mundiales en educación tienden a la inclusión de alumnos con NEET y que no conocemos de ningún estudio en Chile que explore DIF según esta variable en pruebas estandarizadas, resulta relevante estudiar este aspecto. Para ello se analizaron los resultados de una prueba aplicada a niños de 4° a 8° Básico en lenguaje y matemática en un total de más de 100 establecimientos educacionales; esta información se complementó con una encuesta que recopiló información general y específica relativa a las NEET.

3. Método

La investigación presenta un diseño secuencial, definido por Creswell (2008). En una primera etapa, se realizó un análisis estadístico para la determinación de los ítems con DIF seguido por un estudio complementario que indagó la forma en que se diagnostican, reportan y operacionalizan las NEET.

3.1. Análisis estadístico

3.1.1 Participantes

Alumnos de 4° a 8° Básico que respondieron la prueba SEPA de lenguaje y matemática aplicada por MIDE UC en colegios municipales y particulares subvencionados de Chile en 2012. La base de datos para el análisis se segmentó construyéndose muestras que varían entre 2452 y 2912 dependiendo del nivel de enseñanza y asignatura. Las muestras contienen alrededor de 3,4% de niños con NEE.

3.1.2. Instrumento

La prueba SEPA (Sistema de Evaluación Para el Aprendizaje) consiste en un conjunto de pruebas estandarizadas orientadas a evaluar los aprendizajes de los alumnos desde 1° Básico a III Medio en lenguaje y matemática a través de su trayectoria escolar (Pontificia Universidad Católica de Chile, 2014). Es confeccionada por el centro de medición MIDE UC, con base en el marco curricular nacional, y aporta información a docentes y directivos respecto del nivel de aprendizaje de los alumnos en el tiempo. Dicha prueba se aplica en colegios de todo Chile y dependiendo del nivel escolar posee 40 (desde 1° a 6° Básico) o 50 ítems (desde 7° en adelante) de selección múltiple. Durante su aplicación los colegios reportan los alumnos con integración, es decir, con NEE. Sin embargo, la información detallada respecto de cómo se diagnostica, categoriza y se evalúa a los alumnos que presentan NEE en los establecimientos se desconoce, puesto que los colegios presentan autonomía tanto para la declaración de la variable como para la implementación de adecuaciones durante su aplicación.

3.1.3. Procedimiento

Los datos de la prueba SEPA tomados en 2012 se segmentaron considerando: establecimientos que reportaron la variable integración, asignatura (lenguaje o matemática), nivel (4°, 5°, 6°, 7° y 8°) y dependencia (municipales y particulares subvencionados). El estudio consideró solo los niveles de 4° a 8° Básico de colegios municipales y particulares subvencionados que reportaron tener alumnos con NEET debido a que (1) los colegios particulares pagados presentan en promedio un mayor rendimiento que las otras dependencias y una baja presencia de niños con NEET, (2) los niños de 1° a 3° Básico podrían presentar dificultades en las habilidades lectoras propias de su etapa de desarrollo, y (3) el número de alumnos con NEET reportados por nivel disminuye en la medida que aumenta el nivel escolar.

3.1.4. Análisis de datos

La exploración del ajuste de los datos a un modelo unidimensional versus un modelo multidimensional, las diferencias promedio entre los niños con y sin NEET y la evaluación de ítems con DIF se realizaron utilizando el modelo de la Teoría de Respuesta al Ítem (TRI) de un parámetro, dado que los datos se ajustaron mejor a un modelo unidimensional. Este modelo se implementó para cada nivel (4° a 8° Básico) y para cada asignatura (lenguaje y matemática).

El modelo 1PL usado permite: 1) proporcionar mediciones de las variables educativas independientes del instrumento utilizado, y 2) disponer de instrumentos de medida cuyas propiedades no dependen de los objetos medidos (Muñiz, 1997, p. 18). Para Muñiz (1997), los dos supuestos involucrados en la TRI son la curva característica de los ítems (CCI) y la unidimensionalidad.

La CCI muestra que, a medida que aumenta el nivel de aptitud del individuo, se incrementa la probabilidad de acertar correctamente un ítem, lo que es descrito por una función matemática que relaciona la habilidad de los sujetos evaluados y la probabilidad de acertar correctamente un ítem. Dicha función queda definida por tres parámetros: dificultad (b), discriminación (a) y posibilidad de respuesta al azar (c) (Muñiz, 2010).

En un modelo 1PL, la probabilidad de respuesta correcta a un determinado ítem estará determinada por la habilidad del sujeto y la dificultad del ítem (Bond y Fox, 2001). La CCI para un modelo 1PL está dada por la ecuación (1), donde $P_i(\theta)$ es la probabilidad de que aleatoriamente un examinado con habilidad θ elija la respuesta correcta en el ítem i , b_i = parámetro dificultad del ítem i , I = número de ítems en el test y e = base del logaritmo natural cuyo valor aproximado es 2,718.

$$P_i(\theta) = e^{(\theta-b_i)} / [1 + e^{(\theta-b_i)}] \quad I = 1, 2, \dots, I \quad (1)$$

Para el análisis del ajuste de los datos al modelo, la determinación de las diferencias entre grupos y el análisis de ítems con DIF, se utilizó el software Conquest. La forma general del modelo de respuesta al ítem provisto por Conquest es el Modelo Logístico Multinomial Multidimensional de Coeficientes Aleatorios (MRCML) descrito por Adams, Wilson y Wang (1997). El MRCML es una extensión de los modelos que utilizan la TRI de tipo Rasch o 1PL, donde la probabilidad de un sujeto de responder correctamente es función de su capacidad y la dificultad de los ítems. Conquest estima la dificultad de los ítems controlando por el efecto de la variable NEET y el nivel de habilidad (Adams et al., 1997; Wu, Adams y Wilson, 1998).

3.1.5. Estadísticos para evaluar el ajuste del modelo, la diferencia en habilidad entre grupos y la significancia estadística de DIF

Para el análisis de las pruebas se exploró, en primer lugar, el grado de ajuste de los datos al modelo unidimensional versus un modelo multidimensional de cinco dimensiones en lenguaje y cuatro en matemática definidos de acuerdo a las tablas de especificaciones proporcionadas por SEPA.

El ajuste del modelo se examinó utilizando los criterios de información AIC (Akaike) y BIC (Schwarz), dados por las fórmulas (2) y (3), respectivamente, donde k es el número de parámetros estimados en el modelo, n el tamaño de la muestra y L corresponde al valor de la devianza (o varianza) entregadas por el programa.

$$AIC = -2 \log(L) + 2(k + 1) \quad (2)$$

$$BIC = -2 \log(L) + \ln(n)(k + 1) \quad (3)$$

Posteriormente, se estimaron las diferencias promedio entre los grupos con y sin integración según la ecuación (4):

$$\text{Logit}\{P_{ig}(\theta^*)\} = \theta^* - \delta_i + \Delta_g + \gamma_i G \quad (4)$$

Donde $P_{ig}(\theta^*)$ se refiere a la probabilidad de respuesta correcta de una persona perteneciente al grupo g en el ítem i ; θ^* representa la distribución de la capacidad de la persona examinada; δ_i es el parámetro de dificultad del ítem i ; γ_i se refiere a la presencia de DIF en el ítem i ; G indica un grupo de referencia o grupo focal que se comparan entre sí, $G = 0$ si $g = R$ (grupo de referencia), $G = 1$ si $g = F$ (grupo focal); y Δ_g es la diferencia entre la capacidad media de personas en el grupo R y F (Santelices y Wilson, 2012). Para evaluar la significación estadística de la diferencia promedio entre grupos, se prueba la hipótesis nula H_0 : no existe diferencia en la habilidad promedio de los grupos. La diferencia entre grupos distribuye como χ^2 con 1 grado de libertad. Cuando χ^2 ($p < .05$) es significativo, implica que los efectos diferenciales están presentes (Santelices y Wilson, 2012).

Para el análisis de DIF se exploró si existe diferencia en la probabilidad de respuesta correcta a un ítem i , en los grupos con y sin NEET, para un determinado grado de habilidad con el estimador de Wald y con el estadístico *t de student*. El estadístico de Wald tiene una distribución χ^2 con un grado de libertad ($p < .05$). En la prueba de Wald, la estimación de máxima verosimilitud del parámetro de interés θ , es decir $\hat{\theta}$, se compara con el valor propuesto θ_0 , bajo el supuesto de que la diferencia entre ambos será aproximadamente una distribución normal (ver ecuación 6). Típicamente, el cuadrado de la diferencia se compara con una distribución de chi-cuadrado. En el caso univariado estará dado por la ecuación (5):

$$(\hat{\theta} - \theta_0)^2 / \text{Var}(\hat{\theta}) \quad (5)$$

Alternativamente, la diferencia puede ser comparada con una distribución normal. En este caso el resultado es:

$$(\hat{\theta} - \theta_0) / \text{se}(\hat{\theta}) \quad (6)$$

Donde $\text{se}(\hat{\theta})$ es el error estándar de la estimación de máxima verosimilitud.

Dado que el estimador *t-student* involucra supuestos más flexibles que el estimador de Wald, tales como que el tamaño de muestra es grande (asumiéndose distribución normal) y que las diferencias de varianzas entre grupos son desconocidas, se evaluaron los ítems con DIF significativo de acuerdo a la ecuación 6 y $p < .05$.

Para los ítems que presentan DIF significativo se realizó un estudio del tamaño del DIF según Paek (2002) donde se clasifica como DIF imperceptible (A) si el valor estimado es menor que 0.213; presentará DIF moderado (B) si el valor está entre 0.213 y 0.319 y tendrá DIF grande o severo (C) si el valor estimado de DIF es mayor a 0.319 (Santelices y Wilson, 2012).

3.2. Análisis complementario en base a la encuesta sobre NEE

Dados los resultados inesperados obtenidos, se consideró necesario indagar cómo se diagnostica, categoriza y operacionaliza la variable Integración en los establecimientos involucrados en el estudio, con el fin de explorar algunas hipótesis que explicarían dichos resultados.

3.2.1. Participantes

Se encuestó a los coordinadores PIE y/o jefes de UTP de todos los establecimientos disponibles en la base de datos SEPA que aceptaron participar en el estudio pertenecientes a tres comunas populosas de Santiago (Maipú, Macul y Puente Alto). La ciudad de Santiago concentra la mayor parte de la población del país y estas comunas representan al 73% de los colegios de Chile (ADIMARK, 2007).

3.2.2. Instrumento

Se construyó una encuesta de 15 preguntas de respuesta cerrada que fue mejorada según las sugerencias planteadas por académicos de la universidad, un coordinador del proyecto de integración y profesionales de SEPA. Se piloteó el instrumento en dos establecimientos, uno municipal y uno subvencionado, ubicados en las mismas comunas donde se aplicó la encuesta.

3.2.3. Procedimiento

Se contactó telefónicamente a los colegios involucrados en el estudio y se coordinó la aplicación de la entrevista. En ella, cada encuestado firmó un consentimiento escrito y accedió al registro de audio.

3.2.3. Análisis de datos

El análisis de la información recogida se realizó mediante tablas de frecuencia de las respuestas, lo que permitió visualizar similitudes y diferencias.

4. Resultados

4.1. Análisis de DIF

El grado de ajuste de los datos a un modelo unidimensional y multidimensional se evaluó obteniéndose los resultados presentados en la Tabla 1, donde se compara el valor obtenido para L, AIC y BIC (ver ecuaciones 2 y 3).

Si se utiliza solo el valor de la devianza (L_{uni} versus L_{multi}), para el sector de lenguaje, los datos se ajustan mejor a un modelo unidimensional, mientras que para matemática, los datos se ajustan mejor a un modelo multidimensional. Sin embargo, al incorporar

en la evaluación el tamaño de la muestra (n) y el número de parámetros estimados en cada análisis (k), calculando AIC y BIC para todos los niveles, los datos se ajustan mejor a un modelo unidimensional, dado que un valor menor de L, AIC o BIC implica un mejor ajuste al modelo (Santelices y Wilson, 2012).

Tabla 1. Comparación de resultados ajuste a modelo unidimensional y multidimensional usando L, AIC y BIC

PRUEBA	K UNI	L UNI	AIC UNI	BIC UNI	K MULTI	N	L MULTI	AIC MULTI	BIC MULTI
4L	41	139992.5	73.7	323.9	56	2858	140415.8	103.7	443.3
5L	41	140472.1	73.7	324.5	56	2899	141193.2	103.7	444.1
6L	41	123886.5	73.8	318.6	56	2508	125737.9	103.8	436.0
7L	51	174372.3	93.5	402.5	66	2812	176619.9	123.5	521.6
8L	51	147999.9	93.7	395.5	66	2452	148844.2	123.7	512.6
4M	41	133202.0	73.8	322.3	57	2743	133020.5	105.8	448.9
5M	41	143383.6	73.7	324.7	57	2912	144581.3	105.7	452.3
6M	41	130744.8	73.8	320.3	57	2620	130587.2	105.8	446.3
7M	51	176747.8	93.5	403.0	67	2840	176379.9	125.5	530.2
8M	51	162944.4	93.6	398.3	67	2593	162474.3	125.6	524.1

Fuente: Elaboración propia con base en la información reportada por el análisis Conquest.

4.2. Diferencia de habilidad entre grupos

Según los resultados obtenidos con el programa, para las pruebas de lenguaje y matemática aplicadas en los niveles 4°, 5°, 6°, 7° y 8° no existe diferencia significativa en la habilidad promedio entre grupos, puesto que χ^2 en todos los casos es menor que 3.841.

4.2.1. Análisis de ítems que presentan DIF

Al aplicar el criterio de significancia estadística utilizando el estadístico de Wald y *t-student* para las pruebas aplicadas, se obtiene lo reportado en la Tabla 2. Según el estadístico de Wald, no existen ítems con DIF significativo en ninguna de las pruebas aplicadas. Los resultados arrojados por el estadístico *t* en las pruebas de lenguaje aplicadas en 4° y 7° no presentan DIF, mientras que en 5°, los ítems I₃₆ e I₃₇ (5%) favorecen al grupo con integración. En 6°, I₃ perjudica a los alumnos con integración mientras que I₂₇ los favorece. En 8°, 5 ítems presentan DIF (10%); I₁, I₂ e I₇ perjudican a los niños con integración mientras que I₂₉ e I₃₃ los favorecen.

Para la prueba de matemática aplicada, en 5°, 6° y 7° no se evidencia DIF; sin embargo, en 4° los ítems I₁₆, I₂₁ e I₃₅ (7,5% del total) favorecen a los niños con integración. En 8°, se encontraron 4 ítems con DIF (8%); I₄, I₈, I₂₀ perjudicarían a los niños con integración mientras que I₂₅ los favorecería.

4.2.2. Tamaño de DIF

El tamaño del efecto se estableció según lo descrito por Paek en 2002; los resultados obtenidos se muestran en la Tabla 3.

En la prueba de lenguaje de 5° Básico, los ítems con DIF presentan un efecto moderado a favor del grupo con integración. En 6°, I₂₇ presenta un efecto grande a favor del grupo integrado, mientras que I₃ tiene un efecto moderado en contra de ese grupo. En 8° I₁ e I₃₃ presentan un efecto moderado de tendencia mixta, mientras que I₂, I₇ e I₂₉ poseen un efecto grande, también de tendencia mixta.

En la prueba de 4° matemática, I₃₅ presenta un tamaño del efecto imperceptible, I₁₆ muestra un efecto moderado e I₂₁ un efecto grande, todos a favor del grupo integrado. En 8°, I₄ e I₈ tienen un efecto moderado en contra del grupo con integración mientras que I₂₀ e I₂₅ presentan un efecto severo de tendencia mixta.

Tabla 2. Análisis de DIF en prueba SEPA lenguaje y matemática

PRUEBA	Nº DEL ÍTEM	ESTADÍSTICO CALCULADO	WALD (>3,84)	T-STUDENT (>1,96)
4M	16	2.37	no	si
	21	3.27	no	si
	35	1.99	no	si
5L	36	2.08	no	si
	37	2.59	no	si
6L	3	-2.12	no	si
	27	2.34	no	si
8L	1	-1.96	no	si
	2	-2.8	no	si
	7	-2.39	no	si
	29	2.62	no	si
	33	2.20	no	si
8M	4	-2.18	no	si
	8	-2.05	no	si
	20	-2.27	no	si
	25	2.67	no	si

Fuente: Elaboración propia con base en la información reportada por el análisis Conquest.

Tabla 3. Tamaño del efecto para los ítems con DIF significativo

PRUEBA	Nº DEL ÍTEM	ESTADÍSTICO T CALCULADO	DIF ESTIMADO	EFFECTO
5L	36	2.08	0.260	B
	37	2.59	0.280	B
6L	3	-2.12	-0.253	B
	27	2.34	0.356	C
8L	1	-1.96	-0.292	B
	2	-2.80	-0.520	C
	7	-2.39	-0.356	C
	29	2.62	0.382	C
	33	2.20	0.310	B
4M	16	2.37	0.239	B
	21	3.27	0.334	C
	35	1.99	0.203	A
8M	4	-2.18	-0.291	B
	8	-2.05	-0.297	B
	20	-2.27	-0.388	C
	25	2.67	0.355	C

Fuente: Elaboración propia con base en la información reportada por el análisis Conquest.

4.2.3. Encuesta

Los datos recopilados mediante la encuesta se observan en la Tabla 4.

Tabla 4. Resultados encuesta aplicada a establecimientos que reportaron integración en SEPA

COLEGIO	1	2	3	4	5	6
Dependencia	mun	mun	mun	ps	ps	ps
PIE	si	si	si	no	no	si
tipo NEE (NEEP, NEET)	ambas	ambas	ambas	neet	ambas	ambas
Responsable asignación	pie	utp-pie	pie	psic-orient	psic-orient	pie-coor.ped
TDA	si	si	si	si	si	si
TEL	si	si	si	no	si	si
DEA	si	si	si	si	si	si
FIL	si	si	si	si	si	si
Diag. pruebas std	si	si	si	no	no	si
pruebas (SEPA)	si	si	si	no	no	si
informe psicólogo	si	si	si	si	si	si
informe neurólogo	si	si	si	si	si	si
Informe fonoaudiólogo	si	si	si	no	no	si
anamnesis	si	si	si	si	no	si
profesor	si	si	si	si	no	si
N° integrados promedio	15	15	10	5	3	10
NEET frec (TDA)	si	no	no	si	si	si
NEET frec (DEA)	si	si	no	no	no	si
NEET frec (FIL)	si	no	si	no	no	no
REPORTA > 5	no	no	no	no	no	no
Informe SEPA-MIN igual	si	si	si	si	si	si
Adecuaciones	no	no	no	no	no	no

Fuente: Elaboración propia con base en la información reportada por la encuesta.

Notas: MUN = municipal; PS = particular subvencionado; PIE = proyecto integración escolar; UTP = Unidad Técnica Pedagógica; NEEP = necesidades educativas especiales permanentes = NEET = necesidades educativas especiales transitorias; TDA = trastorno déficit atencional; TEL = trastorno específico del lenguaje; DEA = dificultades específicas del aprendizaje; FIL = coeficiente intelectual rango límite; psic= psicólogo; orient= orientador; coor.ped = coordinador pedagógico.

4.3. Análisis de datos recopilados con la encuesta

Los resultados obtenidos a través de la encuesta evidencian que:

En 5 de los 6 los colegios donde se aplicaron las encuestas existen alumnos con NEEP y NEET. Solo en uno (subvencionado) se halló que no existen alumnos con NEEP, sino solo alumnos con NEE transitorias. En todos los establecimientos encuestados existen alumnos con todas las NEET descritas en los antecedentes: TDA, DEA y CI límite (6/6) y TEL en 5/6. En los colegios municipales se comprueba la existencia de alumnos con TDA, DEA, CI límite y TEL (3/3), algo similar ocurre en la dependencia particular subvencionada, donde TDA, DEA y CI límite se encuentra en 3/3 y TEL en 2/3. La NEET más frecuente en los colegios municipales es DEA o FIL con 2/3. En cambio, los colegios particulares subvencionados indican que es TDA (3/3).

La forma de diagnosticar las NEET se relaciona con la existencia de un proyecto de integración (PIE) en el establecimiento. En todos los establecimientos municipales (3/3) y en el único establecimiento subvencionado que cuenta con PIE (1/3), el diagnóstico de NEE es realizado por el equipo de profesionales que conforman el PIE (psicopedagogo, fonoaudiólogo, psicólogo, etc.). En los colegios del sector subvencionado que no cuentan con PIE, el diagnóstico es realizado solo por un psicólogo externo y el orientador del colegio. El Ministerio de Educación de Chile define la documentación requerida para la evaluación de los alumnos con NEET en establecimientos con PIE. Esta incluye: autorización de los padres, diagnóstico con

pruebas estandarizadas, informe psicopedagógico, informe neurológico, informe de fonoaudiólogo, anamnesis, entre otros. En los colegios particulares subvencionados que no cuentan con PIE, la evaluación se basa solo en el informe de un psicólogo y de un neurólogo, incluyendo además en 1/3 colegios anamnesis y en 2/3 la recomendación de un profesor.

El total de alumnos integrados que cada establecimiento reporta oficialmente a SEPA se relaciona con el hecho de tener o no tener PIE. Se observa que los colegios con PIE solo reportan el número máximo de alumnos con NEET que permite la ley ($N = 5$), mientras que los que no tienen PIE reportan un número de alumnos menor o igual a 5. Este fenómeno ocurre de forma indiferente de si son establecimientos municipalizados o subvencionados. Es decir, el número de alumnos integrados reportados a los administradores de la prueba SEPA en ninguno de los 6 colegios supera lo permitido por la ley (5-7). Sin embargo, en todos ellos habría entre 5 y 10 alumnos integrados por curso adicionales, que no habrían sido reportados a SEPA pero si son mencionados en la encuesta. Por otro lado, en los colegios particulares subvencionados sin PIE (2/3), la cantidad de niños con NEET reportada a SEPA coincide con el número de alumnos con NEET reportados en la encuesta.

Todos los establecimientos reportaron que no realizaron adecuaciones durante la aplicación de la prueba.

5. Discusión

En esta investigación se aborda las interrogantes de si el puntaje obtenido en estas pruebas es reflejo del grado de habilidad que el alumno posee y de si está libre de sesgo, es decir, es independiente de la existencia de necesidades educativas especiales transitorias (NEET). Los análisis de DIF realizados en las pruebas de lenguaje y matemática aplicada en los niveles desde 4° a 8° muestran que los ítems con DIF encontrados en 5°, 6° y 8° (en lenguaje) y 4° y 8° (en matemática) no superan el 10%, lo que coincidiría con los resultados de Moreira (2008), quien reportó un 18% de ítems con DIF según TDAH, y con Cho et al. (2012), quienes reportaron un 21% de ítems con DIF.

Los ítems con DIF encontrados en nuestro estudio tanto para la prueba de lenguaje como de matemática son de tamaño moderado o grande de tendencia mixta, es decir, en algunos casos favorecen a los niños con NEET y en otros los perjudican. Esto implica que no existen pruebas suficientes para declarar la existencia de sesgo (favorecer a un grupo sobre el otro) en los resultados de la prueba SEPA, lo que aporta alguna evidencia respecto de la validez de los resultados de dicha prueba.

Además, los resultados sugieren que no existe diferencia en la habilidad promedio entre los grupos estudiados, lo que concuerda con los estudios reportados por Gregg (2007) y por Cho et al. (2012). Sin embargo, no coinciden con lo reportado por Kettler et al. (2005) y Mandinach et al. (2005), puesto que dichos autores sostienen que solo si se aplican adecuaciones al grupo de niños integrados no existirían diferencias entre grupos. Con el fin de comprobar si los resultados obtenidos reflejan similitudes en la habilidad promedio de los grupos o debilidades de la información analizada, se plantearon tres posibles hipótesis. Algunas de ellas fueron exploradas con la encuesta.

La primera hipótesis consideró que los colegios que reportaron integración podrían estar realizando adecuaciones durante la aplicación de la prueba, lo que explicaría un

mejor rendimiento de los niños integrados y por ello no se observarían diferencias en el puntaje promedio entre grupos.

Una segunda hipótesis exploró la posibilidad de que los establecimientos que reportaron integración estuviesen reportando solo una parte de los alumnos, quedando el grupo sin integración con alumnos que sí presentan NEET. Esto indicaría que el reporte de alumnos con NEET por sala estaría condicionado al máximo permitido por la ley (7). En este caso, la comparación estadística entre grupos con NEET y sin NEET perdería algo de validez, ya que el grupo teóricamente sin NEET sería en realidad “híbrido”.

Una tercera posibilidad consideró que los colegios que reportaron integración hubiesen logrado que las dificultades de los alumnos con NEET fuesen superadas con el trabajo de los especialistas, lo que habría resultado en una nivelación de la habilidad promedio entre ambos grupos.

A la luz de los datos reportados por la encuesta, es posible decir que en los colegios que cuentan con PIE, el número de alumnos integrados en aula supera lo descrito por la ley. Esto implica que existió un número de alumnos integrados que no fueron reportados como tales, originándose por una parte un grupo “integrado” de menor tamaño al real y un grupo “sin integración” que contenía en realidad alumnos integrados. Este hallazgo de alguna manera invalida el análisis de DIF realizado en los colegios que, habiendo declarado alumnos integrados a SEPA, presentan también casos de niños con NEET no declarados. Consiguientemente, no es posible inferir la existencia de sesgo en la prueba ya que el comportamiento diferencial de los ítems depende fuertemente de las características de los grupos comparados.

De acuerdo con los resultados de la encuesta, se desestimó que los colegios estuvieran realizando adecuaciones al aplicar la prueba, por lo que la similitud en la habilidad promedio observada no sería debido a esta causa.

Por medio de la encuesta no solo se indagó información relativa a las hipótesis anteriormente mencionadas, sino que se exploraron otros aspectos entre los que podemos mencionar los siguientes:

En primer lugar, los datos recopilados a través de la encuesta permiten corroborar la existencia de NEET y NEEP en las aulas, comprobando que la integración es un hecho patente en los establecimientos del país. En segundo lugar, se evaluó lo descrito por el Mineduc respecto del porcentaje de cada NEET según dependencia en el país. Para el sector municipal, se evidencia que las NEET más frecuentes son DEA y FIL, lo que coincide con los datos disponibles, mientras que en el sector particular subvencionado existen tanto TDA como DEA, pero solo uno de los colegios subvencionados encuestados señala que en su comunidad la NEET más frecuente es DEA, lo que coincidió parcialmente con los datos disponibles, ya que los datos reflejan que la NEET más frecuente es TDA (3/3).

Los datos recolectados respecto de la documentación requerida para diagnosticar a un alumno con NEET y del protocolo de diagnóstico para dichos alumnos descrita por los establecimientos municipales y subvencionados que cuentan con PIE encuestados coinciden íntegramente con lo establecido por el Mineduc. Sin embargo, en los colegios particulares subvencionados que no han implementado PIE existe mayor libertad en cuanto a quién realiza la asignación y con base en qué tipo de documentación. En ellos, se constató que solo se requiere de informe psicopedagógico y/o neurológico.

Cabe destacar que los resultados reportados por el análisis complementario tienen una generalizabilidad limitada debido principalmente al tamaño de la muestra y a que la voluntariedad en la participación puede haber introducido sesgos en las respuestas. Sin embargo, es importante considerar que ella incluye representantes de ambos tipos de dependencia administrativa (municipal y subvencionada) y considera los niveles socioeconómicos más frecuentemente observados en las escuelas de Chile (73% del país), lo que entrega valor a los resultados.

6. Conclusión

A partir del análisis estadístico, se observa que los ítems que presentan DIF en las pruebas de matemática y lenguaje aplicadas no superan el 10% y tienen una tendencia mixta, beneficiando a veces al grupo focal y otras veces al de referencia. Los resultados obtenidos evidencian, sin embargo, que no existiría diferencia en la habilidad promedio entre los grupos estudiados. Dicho resultado coincide solo parcialmente con los resultados de investigaciones anteriores, por lo que se exploró, mediante una encuesta, el posible uso de adecuaciones y la forma en que las escuelas estaban reportando las NEE de sus alumnos.

A partir de los datos recogidos con la encuesta fue posible constatar que los establecimientos no habían realizado ningún tipo de adecuación durante la aplicación de la prueba, por lo que la similitud estadística en la habilidad promedio estimada para el grupo con NEET y sin NEET no sería producto de ello. Aun más, los datos sustentan la hipótesis de que el grupo de referencia utilizado en el análisis DIF sería un grupo de referencia (sin integración) híbrido, ya que los colegios que presentan PIE señalan que el número de alumnos reportados a SEPA coincide con el máximo permitido por ley y que habrían existido alumnos integrados no reportados como tales. A la luz de lo anterior, no es posible inferir respecto de las posibles diferencias en la habilidad promedio entre los grupos de la prueba SEPA según NEET aplicada en lenguaje y matemática en 2012 para los establecimientos municipales y subvencionados con PIE que reportaron menor número de alumnos que los que existían en realidad en aula. Adicionalmente, aunque los ítems con DIF no superan el 10%, no existe claridad en la verdadera causa de estos hallazgos, pudiendo ser por diferencias estadísticas en el comportamiento del ítem o por deficiencias en la asignación de grupos.

Una futura investigación debería considerar un análisis de la existencia de DIF según NEET asociado a encuestas y trabajo de campo que permita asegurar que el grupo sin integración no es híbrido, es decir, que lo declarado por los establecimientos en la prueba aplicada coincide íntegramente con lo evidenciado en aula y no solo es coincidente con lo que se espera desde lo legal.

Referencias

- Adams, R., Wilson, M. y Wang, W. (1997). The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement*, 21(1), 1-23. doi:10.1177/0146621697211001.
- ADIMARK. (2007). *Mapa socioeconómico de Chile: Nivel socioeconómico de los hogares del país basado en datos del Censo*. Recuperado de http://www.adimark.cl/medios/estudios/Mapa_Socioeconomico_de_Chile.pdf
- Baena, G. (1998). *Técnicas de investigación*. Ciudad de México: Editores Mexicanos Unidos.

- Bond, T. y Fox, C. (2001). *Applying the Rasch model*. Mahwah, NJ: Lawrence Erlbaum.
- Cahalan-Laitusis, C. (2010). Examining the impact of audio presentation on tests of reading comprehension. *Applied Measurement in Education*, 23(2), 153-167. doi:10.1080/08957341003673815.
- Cawthon, S. (2010). Examining student factors in sources of setting accommodation DIF. *Educational and Psychological Measurement*, 74(5), 759-794. doi:10.1107/0013164413514053.
- Cho, H., Lee, J. y Kingston, N. (2012). Examining the effectiveness of test accommodation using DIF and a mixture IRT model. *Applied Measurement in Education*, 25(4), 281-304. doi:10.1080/08957347.2012.714682.
- Cohen, A., Gregg, N. y Deng, M. (2005). The role of extended time and item content on a High-Stakes Mathematics Test. *Learning Disabilities Research & Practice*, 20(4), 225-233. doi:10.1111/j.1540-5826.2005.00138.x.
- CONEVAL. (2012). *Informe de pobreza y evaluación en el estado de Sonora*. Recuperado de <http://desarrollosocial.guanajuato.gob.mx/coneval/informe-sonora.pdf>
- CONEVAL. (2012). *Porcentaje de población en situación de pobreza extrema según entidad federativa*. Recuperado de http://www.coneval.gob.mx/SalaPrensa/Documents/Comunicado005_Medicion_pobreza_2014.pdf
- Creswell, J. (2008). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (3ª ed.). California, CA: Pearson.
- Delval, J. (1992). *Aprender a aprender. El desarrollo de la capacidad de pensar*. Madrid: Alhambra Longman.
- Elosúa, P. (2003). Sobre la validez de los test. *Psicothema*, 15(2), 315-321.
- Finch, W. y Finch, M. (2013). Investigation of specific learning disability and testing accommodations based differential item functioning using a multilevel multidimensional mixture item response theory model. *Educational and Psychological Measurement*, 73(6), 973-993.
- Gómez-Benito, J., Hidalgo, M. y Guilera, G. (2010). El sesgo de los instrumentos de medición. *Tests justos. Papeles del Psicólogo*, 31(1), 75-84.
- Gregg, N. (2007). Underserved and unprepared: Postsecondary learning disabilities. *Learning Disabilities Research & Practice*, 22(4), 219-228. doi:10.1111/j.1540-5826.2007.00250.x.
- Hen, M. y Goroshit, M. (2014). Academic procrastination, emotional intelligence, academic self-efficacy, and GPA: A comparison between students with and without learning disabilities. *Journal of Learning Disabilities*, 47(2), 116-124. doi:10.1177/0022219412439325.
- Idan, O. y Margalit, M. (2014). Socioemotional self-perceptions, family climate, and hopeful thinking among students with learning disabilities and typically achieving students from the same classes. *Journal of Learning Disabilities*, 47(2), 136-152. doi:10.1177/0022219412439608.
- INEGI. (2014, 27 de octubre). *Consulta interactiva de datos*. Recuperado de <http://www.inegi.org.mx/sistemas/olap/proyectos/bd/consulta.asp?p=17118yc=27769ys=est>
- Infante, M., Matus, C. y Vizcarra, R. (2010). Desafíos a la formación docente: Inclusión educativa. *Estudios Pedagógicos*, 36(1), 287-297. doi:10.4067/s0718-07052010000100016.

- Kettler, R. J., Niebling, B. C, Mroch, A. A., Feldman, E. S., Newell, M. L., Kratochwill, T. R. ... y Bolt, D. M. (2005). Effects of testing accommodations on Math and Reading scores: An experimental analysis of the performance of students with and without disabilities. *Assessment for Effective Intervention*, 31(1), 37-48. doi:10.1177/073724770503100104.
- Lin, P. y Lin, Y. (2014). Examining student factors in sources of setting accommodation DIF. *Educational and Psychological Measurement* 74(5), 759-794. doi:10.1177/0013164413514053.
- Mandinach, E. B., Bridgeman, B., Cahalan-Laitusis, C. y Trapani, C. (2005). The impact of extended time on SAT test performance. *The College Board*, 2, 1-35.
- Meltzer, Y. (2006). תמיכה מרכז חי-תל האקדמית במכללה למידה לקויות עם לסטודנטים . [A support center for learning disabled students in Tel-Hai Academic College]. *Issues in Special Education & Rehabilitation*, 21, 98-101.
- Ministerio de Educación de Chile. (2009). *Decreto N° 170/09. Normativa Educación Especial*. Recuperado de <http://www.mineduc.cl/usuarios/edu.especial/doc/201304231500550.DEC200900170.pdf>
- Moreira, T. E. (2008). Funcionamiento diferencial del ítem en pruebas de matemática para educación media. *Actualidades en Psicología*, 22(109), 91-113. doi:10.15517/ap.v22i109.16.
- Muñiz, J. (1997). *Introducción a la teoría de respuesta a los ítems*. Madrid: Pirámide.
- Muñiz, J. (2010). Las teorías de los test: Teoría clásica y teoría de respuesta a los ítems. *Papeles del Psicólogo*, 31(1), 57-66.
- Navas, M. J. (2012). La medición en el ámbito educativo. *Psicología Educativa*, 18(1), 15-28. doi:10.5093/ed2012a2.
- Organización de las Naciones Unidas. (2008). *Declaración Universal de los Derechos Humanos*. Recuperada de en <http://www.temoa.info/es/node/19618>
- Paek, I. (2002). *Investigation of differential item functioning: Comparisons among approaches, and extension to a multidimensional context*. (Tesis doctoral no publicada). Universidad de California.
- Pontificia Universidad Católica de Chile. (2014, 23 de septiembre). *Sistema de Evaluación de Progreso del Aprendizaje: Pruebas SEPA*. Recuperado de <http://www.sepauc.cl/wp-content/uploads/2014/05/20140923-Sistema-Evaluacion.pdf>
- Prieto, G. y Delgado, A. (2003). Análisis de un test mediante el modelo de Rasch. *Psicothema*, 15(1), 94-100.
- Ramos, R. Y. (2001). *Educación integral. Una educación holística para el siglo XXI* (Vol I). Madrid: Desclée de Brouwer.
- Santelices, M. V. y Wilson, M. (2012). On the relationship between differential item functioning and item difficulty: An issue of methods? Item response theory approach to differential item functioning. *Educational and Psychological Measurement*, 72(1), 5-36. doi:10.1177/0013164411412943.
- Sireci, S., Scarpati, S. y Li, S. (2005). Test accommodations for students with disabilities: An analysis of the interaction hypothesis. *Review of Educational Research*, 75(4), 457-490. doi:10.3102/00346543075004457.
- Tenorio, S. (2011). Formación inicial docente y necesidades educativas especiales. *Estudios Pedagógicos*, 37(2), 249-265. doi:10.4067/s0718-07052011000200015.
- Wu, M., Adams, R. J. y Wilson, M. (1998). *ACER-ConQuest*. Hawthorn, Victoria: ACER Press.

El Programa de Evaluación Docente en la UADY. Experiencia de la Implementación

Teaching Assessment Program in the UADY. Implementation experience

Norma Alicia Benois Muñoz*, Fabiola Isabel Gómez Cauich,
Maritza Minelli Briceño Caballero y Jessica Betzabé Zumárraga Ávila

Universidad Autónoma de Yucatán

La evaluación del desempeño docente tiene la finalidad de valorar la calidad de las responsabilidades en la enseñanza, aprendizaje y desarrollo del estudiante, sin embargo, es una tarea difícil de realizar (SEP, 2010). Debido a la necesidad de contar con un proceso de evaluación que contribuya con la mejora de la enseñanza, se creó el Programa de Evaluación Docente en la Universidad Autónoma de Yucatán (UADY). Éste propone promover la reflexión y el mejoramiento de la práctica docente, fortalecer los sistemas de formación y capacitación, así como estimular el interés de los profesores por el desarrollo profesional.

El objetivo de este trabajo es presentar la experiencia de su implementación y los resultados obtenidos en el primer ejercicio. Se realizó un estudio descriptivo con un enfoque mixto, predominantemente cuantitativo (Hesse-Biber, 2010). Participaron estudiantes y profesores del primer y segundo semestre de siete programas educativos. Los profesores obtuvieron resultados positivos en la evaluación de su desempeño, se reportó un avance entre los resultados del primer y segundo semestre y se observó correspondencia entre las respuestas otorgadas por estudiantes y profesores al evaluar el desempeño. De la misma forma, se identificaron las fortalezas y áreas de mejora del desempeño de los profesores.

Palabras clave: Evaluación docente, Educación superior, Estrategias de evaluación.

Teaching assessment is intended the purpose to make judgements regarding the quality of teachers' responsibilities accomplishment related to teaching, learning and students development, yet evaluating it seems to be a difficult task considering that teaching is very a complex and wide work (SEP, 2010). Because of the need of assessing teaching to better teachers' performance, the Autonomous University of Yucatan (UADY) developed "The Teaching Assessment Program" which promotes a reflexive practice of teacher's own performance. Through this program the UADY intend to strengthen the ongoing training system and encourage teachers into professional development.

The aim of this paper is to present the experience of their implementation and the results obtained in the first year. The current article presents a descriptive study with a mixed approach of investigation, quantitative predominantly (Hesse-Biber, 2010). This study presents the opinion of teachers and students of first and second semesters of seven different educational programs regarding their structure and the implementation as well as the teachers' performance. Although it was found that teachers usually obtain good assessment's results, their performance was better assessed in second semester than it was in the first one; both teachers and students coincided with teachers' performance assessment results. It also was found some strong areas and areas for improvement of the seven educational programs that in teachers and students' opinion must be considered to improve teachers' performance.

Keywords: Assessment of teachers, Higher education, Evaluation strategies.

*Contacto: norma.benois@correo.uady.mx

1. Introducción

La evaluación docente es uno de los temas que ha cobrado más relevancia dentro del ámbito educativo, ya que se concibe como:

Un proceso que permite obtener información clara, precisa y confiable sobre las acciones que emprende una institución educativa, dando lugar a la retroalimentación y toma de decisiones que permitan elevar la calidad de la enseñanza y fortalecer la formación docente. (Aguilar, 2009, p. 1)

En sus inicios, la evaluación se centraba en la fiscalización, como un mecanismo de control; sin embargo, la tendencia afirma que debe enfocarse en la mejora de las actividades que se evalúan (García, Loredó, Luna y Rueda, 2008).

En concordancia con lo anterior, el Modelo Educativo para la Formación Integral (MEFI) de la Universidad Autónoma Yucatán (UADY) establece la evaluación al profesorado como un medio para promover su desempeño en términos del “perfil del profesor UADY” y considera que la evaluación debe hacerse desde diferentes perspectivas. A partir de esto, en el 2012 se elabora la propuesta para implementar el Programa de Evaluación Docente, cuyos propósitos son:

1. Promover la práctica reflexiva del docente sobre su propio desempeño y el compromiso frente a la calidad de la enseñanza.
2. Favorecer el mejoramiento de la práctica docente.
3. Contribuir al fortalecimiento de los sistemas de formación y capacitación del profesorado.
4. Estimular el interés de los profesores por el desarrollo profesional.

Para alcanzar tales propósitos, se toman en consideración tres elementos: evaluación por parte del estudiante, portafolio de evidencias y autoevaluación. Para ello, se diseñaron instrumentos con base en las actividades y competencias establecidas en el perfil del profesor UADY. El objetivo de este trabajo es presentar la experiencia de su implementación y los resultados obtenidos en el primer año de ejecución.

2. Fundamentación teórica

2.1. El trabajo docente

El trabajo docente es uno de los temas que ha generado gran interés de las instituciones, gobiernos y académicos de todo el mundo, ya que es reconocida la importancia que tiene el trabajo del profesor en la formación de estudiantes y la producción del conocimiento. Sin embargo, la complejidad de las tareas y los roles que desempeña, así como diversos factores que intervienen en su labor, plantea la necesidad de repensar cómo se concibe al profesor y la atención que debe ponerse en su formación y mejora continua.

En concordancia con esta tendencia, la UADY (2012, p. 48), declara en el Modelo Educativo vigente, que es necesario dimensionar el rol del profesor, para responder a las nuevas demandas que se presentan y establece el perfil deseable del profesor. En este se declara que el profesor es quien tiene la obligación de crear las condiciones de aprendizaje para el desarrollo de competencias; debe realizar actividades como facilitador, tutor, asesor, gestor y evaluador de los aprendizajes. Esto es:

- a) Facilitador, porque planea las experiencias de aprendizaje significativo a partir del reconocimiento e identificación de los problemas del contexto; crea

y recurre a escenarios de aprendizaje reales que favorezcan el desarrollo de las competencias dentro y fuera del aula, y apoya al estudiantado en la construcción del conocimiento.

- b) Tutor, porque guía al estudiante en la toma de decisiones académicas, administrativas, personales y profesionales, favoreciendo la permanencia y conclusión exitosa de su formación profesional y, cuando es necesario, lo transfiere a otro tipo de apoyos.
- c) Asesor, porque apoya al estudiante en relación con dificultades encontradas en el proceso de enseñanza y aprendizaje y orienta su buen desempeño académico en materia de contenidos y áreas disciplinares.
- d) Gestor, porque procura el acceso a escenarios, recursos y contextos académicos a los que el estudiante por sí solo no podría acceder.
- e) Evaluador, porque diseña la evaluación, propone criterios y evidencias de desempeño, emite juicios de valor en diversos momentos y circunstancias, y retroalimenta permanentemente el grado en que el estudiante ha desarrollado las competencias esperadas.

Del mismo modo, se declara que para el diseño de experiencias de aprendizaje, es necesario que el profesor considere las características del estudiantado y su contexto, el área disciplinar, los contenidos, los resultados de aprendizaje esperados, la modalidad de estudio (presencial, no presencial y mixta), la utilidad de las TIC, la disponibilidad de los recursos e infraestructura académica disponible y la posibilidad de acceder a escenarios de aprendizaje fuera del aula. Por tanto, en el marco de este Modelo Educativo, es necesario que los profesores se corresponsabilicen con la institución y esta, a su vez, favorezca el desarrollo de sus competencias que les permita ponerlo en práctica al interior y exterior de las aulas (UADY, 2012).

Como puede observarse, el Modelo Educativo vigente en la UADY otorga gran importancia al profesor, al considerarlo como uno de los actores principales del proceso de enseñanza y aprendizaje y como un participante activo en la formación del estudiante. La declaración de un perfil denota un interés de la institución en su trabajo docente, en su formación y actualización continua como la clave que garantice la calidad de la educación.

Por tanto, para el logro del perfil deseable del profesor es necesario promover la cultura de la evaluación y la promoción de la mejora continua entre el personal docente y en este sentido, la evaluación del desempeño docente surge como una acción institucional para favorecer lo anterior y los resultados obtenidos se consideran un insumo para que el profesor reflexione sobre su práctica.

2.2. La evaluación del desempeño docente

La evaluación del desempeño docente puede ser concebida de diversas formas. Las diferencias pueden estar determinadas por las políticas y el contexto de la institución en la que se lleva a cabo, los objetivos, uso de los resultados y los fines para los cuales fue creada. De acuerdo con Rizo (1999, p. 1), “la evaluación del desempeño docente es un proceso permanente, enmarcado dentro de una concepción de calidad de la educación y que se enfoca en el perfeccionamiento de la docencia en una institución educativa”.

Por su parte, De Miguel (1998) la define como un proceso que permite al profesor detectar sus fortalezas y debilidades, así como los aspectos en los que debe mejorar. Afirma que en toda evaluación es necesario que el profesor se involucre y participe de forma activa para lograr la mejora a largo plazo.

En México, la Secretaría de Educación Pública define que la evaluación docente es:

Un proceso cuya finalidad es emitir juicios de valor sobre la calidad del cumplimiento de responsabilidades docentes en la enseñanza, aprendizaje y desarrollo de los estudiantes, a través de un seguimiento permanente que permita obtener información válida, objetiva y fiable para determinar los avances alcanzados en relación con los propósitos de la enseñanza y propósitos educativos. (SEP, 2010, p. 11)

Actualmente la mayoría de las Instituciones de Educación Superior han puesto en marcha iniciativas para evaluar el desempeño docente. En el contexto latinoamericano, se encuentra el Sistema de Evaluación del Desempeño Profesional Docente en Chile, el cual está orientado al fortalecimiento de la profesión docente. Desde su origen en la década de los noventa ha estado relacionado con diversos programas e iniciativas que buscan dignificar y restituir el carácter profesional de la actividad pedagógica (Bonifaz, 2011, p. 15).

En México, las acciones de evaluación del desempeño docente se identifican fundamentalmente como parte de los programas de compensación salarial, con el Programa de Mejoramiento del Profesorado (PROMEP), cuyo nombre cambió a partir del año 2014 a Programa para el Desarrollo Profesional Docente (PRODEP), el Programa de Fortalecimiento Institucional (PIFI) nombrado a partir de 2013 como Programa de Fortalecimiento de la Calidad de las Instituciones, entre otros. En la mayor parte de las Instituciones de Educación Superior, la evaluación docente es considerada como condición para el acceso a estos programas y también como insumo para la creación e implementación de iniciativas para el desarrollo profesional del profesor (Rueda, 2008).

Lo anterior denota la necesidad de crear más programas enfocados en la mejora y en la reflexión de la práctica del profesor. Para ahondar en este tema, se presentan los resultados de algunas investigaciones enfocadas en la evaluación del desempeño docente en las Instituciones de Educación Superior.

En primera instancia, Cisneros y Aguilar (2010) encontraron que la evaluación es necesaria para que los profesores obtengan retroalimentación de su desempeño docente y pueda proporcionar información valiosa para mejorar el mismo. Sin embargo, existen problemáticas relacionadas con la evaluación que deben tomarse en cuenta, una de ellas es la subjetividad, pues consideran que “la opinión de los estudiantes se ve afectada por su interés en el curso, la personalidad o estilo del profesor y especialmente por sus calificaciones” (Cisneros y Aguilar, 2010, p. 248).

Entre las dificultades identificadas en la investigación de Cisneros y Aguilar (2010) se encuentra el hecho de que, en la mayoría de los casos, no se evalúan todas las competencias de la práctica docente. Además, los profesores consideran que hay inconvenientes con los instrumentos utilizados, dado que algunos reactivos no son claros en su redacción, lo cual puede afectar la interpretación de los estudiantes al responder los cuestionarios. Aunado a lo anterior, se encontró que existe temor entre los docentes en cuanto al uso de los resultados y manifestaron su preocupación, ya que a nivel institucional se da importancia a los resultados para rendir cuentas respecto a indicadores propuestos por agencias externas, como el Padrón de Posgrados del

Consejo Nacional de Ciencia y Tecnología (CONACYT) y los procesos de Acreditación de los Comités Interinstitucionales para la Evaluación de la Educación Superior (CIEES) (Cisneros y Aguilar, 2010).

De la misma forma, Rueda (2008) realizó un trabajo en donde se sintetizan las características más sobresalientes de la evaluación del desempeño docente en las universidades públicas de la Ciudad de México, entre sus conclusiones afirmó:

Debido a la diversidad de ambientes institucionales, modelo educativo, infraestructura y condiciones del proceso de enseñanza aprendizaje, es importante contextualizar la evaluación es decir considerar en el diseño y la puesta en marcha el contexto institucional en que se encuentra. (Rueda, 2008, p. 10)

Por su parte, Cisneros y Aguilar (2010) enfatizaron la importancia de informar a los evaluados acerca del proceso además:

Es necesario que la institución defina y dé a conocer los objetivos de la evaluación y el uso de los resultados. Agregaron que es necesario contar con espacios para la reflexión y apoyar a los docentes a mejorar su práctica. (Cisneros y Aguilar, 2010, p. 249)

Los resultados del estudio de Cisneros y Aguilar (2010) indican que aunque comienza a considerarse el uso de otras fuentes de información en la evaluación de la docencia, tales como evaluación por pares, auto-evaluación y evaluación por supervisores o coordinadores, todavía no se utilizan en forma regular y no se han tomado en consideración las diferencias entre estas fuentes de información, así como sus fortalezas y debilidades cuando se usan para emitir un juicio sobre la calidad del desempeño docente.

Por su parte García, Loredo, Luna y Rueda (2008) mencionan que es necesario concebir la evaluación como un proceso tan complejo como la docencia misma. Asimismo, que la enseñanza puede ser juzgada de manera apropiada solo si se evalúa en el marco de los factores que la determinan. Bajo esta perspectiva se esperaría contar con sistemas acordes con el contexto educativo y características de los profesores.

Como puede observarse en las investigaciones presentadas existen múltiples formas de concebir la evaluación del desempeño docente, se han detectado problemáticas y aspectos a considerar dentro de los procesos de evaluación, por lo tanto, es recomendable que cualquier programa de evaluación docente que esté en proceso de construcción o en sus primeras etapas de implementación, considere esta información para contribuir al éxito y maximizar los beneficios del proceso.

Por tanto, el Programa de Evaluación Docente en la UADY surge con el propósito de contribuir al desarrollo y fortalecimiento de la profesión docente, específicamente al mejoramiento de la labor pedagógica de los profesores, con el fin de favorecer el aseguramiento de aprendizaje de calidad de los estudiantes y aportar información valiosa a los actores educativos. Para su construcción se tomó en consideración el contexto en el que se implementará, la opinión de los actores del proceso de enseñanza y aprendizaje, así como la congruencia con su principal referente: el Modelo Educativo vigente.

Este Programa concibe la evaluación con un carácter en primera instancia formativo y constituye una oportunidad para que el docente se conozca mejor desde el punto de vista profesional, identificando tanto sus fortalezas como los aspectos que puede mejorar. A continuación se detalla la estructura del Proceso.

2.3. Estructura del Proceso de Evaluación

En la Universidad Autónoma de Yucatán, la evaluación del desempeño docente se realiza a través del Programa de Evaluación Docente que se ha implementado en los Programas Educativos de licenciatura, alineados al MEFI desde agosto de 2013. Este programa se integra por tres elementos:

- a) Evaluación por parte del estudiante. Esta evaluación se realiza por medio de un formulario en versión electrónica que se aplica unas semanas previas al término del período escolar a evaluar.
- b) Autoevaluación del profesor. Esta evaluación se realiza por medio de un formulario en versión electrónica alojada en una plataforma denominada “UADY Virtual”, la cual se solicita al profesor que conteste en las últimas semanas del período escolar.
- c) Portafolio de evidencias. Se solicita al profesor la creación de un portafolio electrónico el cual consta de tres partes: introducción, evidencias y reflexión final.

La ponderación de los elementos de evaluación se analizó y se determinó por un grupo de asesores expertos. Se sugirió dar mayor peso a la opinión de los estudiantes y dividir la ponderación de la evaluación por parte del propio profesor en dos elementos: el instrumento de autoevaluación y el portafolio de evidencias. La ponderación resultante fue la mostrada a la tabla 1.

Tabla 1. Ponderación de los elementos de la evaluación

ELEMENTO DE EVALUACIÓN	
Evaluación por parte del estudiante	60%
Autoevaluación del profesor	15%
Portafolio de evidencias	25%

Fuente: Elaboración propia.

Una vez recuperada la información de estos elementos, se obtiene una puntuación final sobre 100 puntos, los cuales permiten ubicar el nivel de desempeño del profesor de acuerdo a la tabla 2.

Tabla 2. Nivel de desempeño de los profesores

RANGO	NIVEL DE DESEMPEÑO
91-100	Excelente
76-91,99	Satisfactorio
61-75,99	Suficiente
1-60,99	Necesita mejorar

Fuente: Elaboración propia.

En virtud de que el Programa de Evaluación Docente en la UADY comenzó su implementación recientemente, resulta necesario documentar sus primeros resultados y compartir la experiencia con el fin de identificar aspectos de mejora que contribuyan al éxito del mismo. De la misma forma, se busca desarrollar una evidencia de su construcción y un punto de partida para futuros estudios relacionados.

3. Metodología

Tipo de estudio y diseño. Se presenta un estudio descriptivo con un enfoque mixto de investigación, predominantemente cuantitativo (Hesse-Biber, 2010).

Participantes. Se presenta la información recopilada en los dos primeros semestres de la implementación del programa de evaluación. Estos son: agosto-diciembre de 2013 y enero-mayo de 2014. Los participantes son estudiantes y profesores del primer y segundo semestre de siete programas en los Campus de Ciencias de la Salud CCS (Facultad de Odontología, Química y Enfermería), Campus de Ciencias Exactas e Ingenierías CCEI (Facultad de Matemáticas) y el Campus de Ciencias Biológicas y Agropecuarias CCBA (Facultad de Veterinaria). El desglose se presenta en la tabla 3.

Tabla 3. Relación de profesores y estudiantes que participaron en el proceso de evaluación

CAMPUS	FACULTAD	PROFESORES	ESTUDIANTES	PROFESORES	ESTUDIANTES
		SEMESTRE AGO-DIC 2013	SEMESTRE AGO-DIC 2013	SEMESTRE ENE-MAY 2014	SEMESTRE ENE-MAY 2014
CCBA	Veterinaria	17	156	20	80
	Enfermería	11	78	14	125
CCS	Odontología	10	83	12	94
	Química	14	102	16	88
CCEI	Matemáticas	9	91	8	59
Total		61	510	70	446

Fuente: Elaboración propia.

3.1. Instrumentos

Para recolectar la información sobre la evaluación se construyeron y validaron dos: el Instrumento de evaluación del profesor por parte de los estudiantes y el Instrumento de autoevaluación del profesor. Además, se solicitó la elaboración de un portafolio de evidencias. Se consideraron estos tres elementos como un primer acercamiento, dado que era necesario conocer la opinión del estudiante como usuario y poder contrastarla con la propia opinión del profesor. Por ello, se realizaron pruebas espejo de estudiantes y profesores. Por último, se consideró la elaboración de un portafolio de evidencias que sirva como un instrumento adicional de autoevaluación y reflexión del propio desempeño docente. Cada elemento se describe a continuación.

3.1.1. Instrumento de evaluación del profesor por parte de los estudiantes

Se compone de 35 ítems organizados en siete secciones (tabla 4). Para obtener el coeficiente de confiabilidad se utilizó el Alpha de Cronbach, del cual se obtuvo un coeficiente total de 0,905. Esta puntuación se validó con la prueba de T de Hotelling con puntaje de 654,874 y $p = 0,000$. En la tabla 4 se muestran los coeficientes obtenidos para cada sección. Cabe señalar que la última, denominada “satisfacción del estudiante”, no entró en el análisis por utilizar otra escala de respuesta.

Tabla 4. Relación de secciones e ítems correspondientes con su coeficiente de confiabilidad

SECCIONES	ÍTEMES	COEFICIENTE DE CONFIABILIDAD
Actividades como facilitador	1-10	0,75
Actividades como evaluador	11-18	0,74
Actividades como asesor	19-24	0,82
Actividades como gestor	25	0,74
Actividades como tutor	26-27	0,73
Actitud del profesor	28-30	0,81
Satisfacción del estudiante	31-35	

Fuente: Elaboración propia.

Para las seis primeras secciones, en el período agosto-diciembre de 2013, se utilizó una escala de uno a cinco puntos, donde: 1 = casi nunca, 2 = ocasionalmente, 3 = algunas veces, 4 = frecuentemente y 5 = casi siempre.

Posterior a la primera aplicación formal del instrumento, se realizó una metaevaluación, en la que se les pidió a los profesores que expresaran su opinión respecto a los instrumentos utilizados. El resultado de esta derivó en la adecuación de la escala utilizada, dado que se sugirió prescindir del punto medio y con ello evitar la tendencia a la respuesta neutral o indiferente. Por tanto, en el período de enero-mayo de 2014, se modificó de uno a cuatro puntos donde: 1 = casi nunca, 2 = algunas veces, 3 = frecuentemente y 4 = casi siempre.

En la última sección de la escala, se utilizó una escala de uno a tres puntos, donde 1 = insatisfecho, 2 = satisfecho y 3 = muy satisfecho.

Por último, se incluyen cuatro preguntas abiertas en las que se consideran los aspectos beneficiosos de la asignatura, las fortalezas y áreas de mejora de los profesores y comentarios adicionales.

3.1.2. Instrumento de autoevaluación del profesor

Se compone de 36 ítems organizados en seis secciones (tabla 5).

Tabla 5. Relación de secciones e ítems correspondientes al instrumento de autoevaluación con su coeficiente de confiabilidad

SECCIONES	ÍTEMS	COEFICIENTE DE CONFIABILIDAD
Actividades como facilitador	1-10	0,83
Actividades como evaluador	11-21	0,83
Actividades como asesor	22-27	0,87
Actividades como gestor	28-30	0,82
Actividades como tutor	31-33	0,79
Actitud del profesor	34-36	0,88

Fuente: Elaboración propia.

Utilizando el mismo procedimiento que para el instrumento de estudiantes, para obtener el coeficiente de confiabilidad se utilizó el Alpha de Cronbach, del cual se obtuvo un coeficiente total de 0,957. Esta puntuación se validó con la prueba de T de Hotelling con puntaje de 456,191 y $p = 0,000$. En la tabla 5 se muestran los coeficientes obtenidos para cada sección.

De igual forma, en el período de agosto-diciembre de 2013, para estas seis secciones se utilizó una escala de uno a cinco puntos, donde: 1 = casi nunca, 2 = ocasionalmente, 3 = algunas veces, 4 = frecuentemente y 5 = casi siempre.

Dado que el resultado de la metaevaluación de la primera aplicación formal del instrumento derivó en un cambio de escala para el instrumento de estudiantes y con la intención de mantener la evaluación espejo, en el período de enero-mayo de 2014 se utilizó una escala de uno a cuatro puntos, donde: 1 = casi nunca, 2 = algunas veces, 3 = frecuentemente, 4 = casi siempre.

Por último, se anexan tres preguntas abiertas en donde se consideran las fortalezas y áreas de mejora de los profesores y comentarios adicionales.

3.1.3. Portafolio de evidencias

El portafolio consta de tres partes: introducción, evidencias y reflexión final.

- a) La introducción se redacta en forma de ensayo, se solicita que se presenten los datos de identificación del profesor y la asignatura a evaluar y, además, se integra la filosofía de la docencia del profesor.
- b) En el apartado de evidencias se anexan materiales y actividades utilizados dentro el proceso de enseñanza y aprendizaje como su planeación didáctica, evaluaciones y actividades de aprendizaje de los estudiantes, en donde se analice la pertinencia de dichas actividades, los criterios de evaluación utilizados para estas, así como las ventajas y desventajas del uso de las actividades en cuestión.
- c) La reflexión final también se redacta a manera de ensayo y se analiza si se alcanzaron las competencias establecidas en la planeación didáctica, las estrategias y factores que favorecieron el desarrollo de las competencias, así como las que obstaculizaron, los cambios propuestos para futuros cursos, además de necesidades de capacitación o áreas de oportunidad en las que se compromete a trabajar para mejorar su desempeño docente.

Procedimiento: La recogida de información se realizó durante dos períodos escolares: agosto-diciembre de 2013 y enero-mayo de 2014. La evaluación se realizó con los profesores con asignaturas alineadas al MEFI y solo participaron las cinco facultades mencionadas con anterioridad. Para realizar la evaluación, al inicio de cada período escolar se solicitó en las facultades la relación de profesores y asignaturas a impartir y se establecieron fechas para la implementación de la evaluación previo al término de cada período escolar. Una vez obtenida la información, se elaboró una base de datos en SPSS para realizar el análisis cuantitativo de los cuestionarios para estudiantes y profesores. Aunado a esto, se realizó la categorización de las respuestas de las preguntas abiertas de cada cuestionario para agregarlas a los reportes de evaluación. Por último, se evaluó el portafolio del profesor por medio de una rúbrica.

Una vez conjuntada la información de cada elemento de evaluación, se generó un reporte por cada asignatura que impartió un profesor. Estos reportes se entregaron a los directivos de las facultades para trabajar la retroalimentación de los profesores.

4. Resultados

Para atender al objetivo de este trabajo, que es presentar la experiencia de su implementación y los resultados obtenidos en el primer año de ejecución, se muestran los resultados más relevantes obtenidos en el ciclo 2013-2014. La información se expone a modo de comparación por período escolar con: a) resultados generales, b) resultados generales por campus, c) resultados generales por facultad, d) comparaciones entre las respuestas de estudiantes y profesores por rol, e) análisis de comentarios de los estudiantes y profesores y, por último, f) resultados recabados en el portafolio del profesor.

4.1. Resultados generales por semestre

Es favorable observar que para el período de enero-mayo de 2014 se incrementó el porcentaje de profesores evaluados con nivel excelente y satisfactorio y no hubo profesores evaluados con el nivel necesita mejorar (gráfico 1).

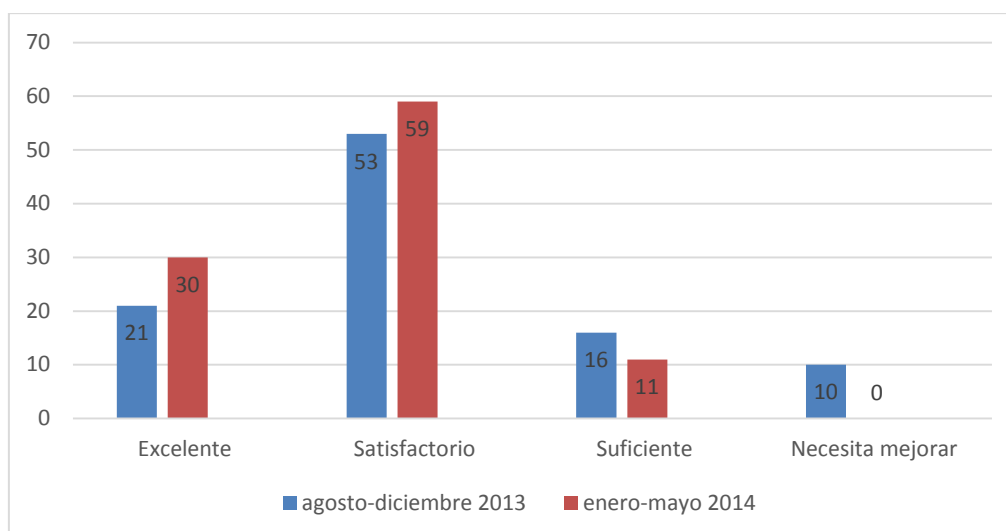


Gráfico 1. Comparación del porcentaje de profesores evaluados de acuerdo a su nivel de desempeño

Fuente: Elaboración propia.

4.2. Resultados por campus: Comparación semestral

Como se observa en la tabla 6, tanto en el Campus de Ciencias de la Salud, como en el de Ciencias Biológicas y Agropecuarias, se incrementó el porcentaje de profesores con nivel de desempeño excelente.

Tabla 6. Comparación del porcentaje de profesores evaluados de acuerdo a su nivel de desempeño por Campus y período de evaluación

NIVEL DE DESEMPEÑO POR CAMPUS	EXCELENTE		SATISFACTORIO		SUFICIENTE		NECESITA MEJORAR	
	1	2	1	2	1	2	1	2
Período*								
CCS	20	36	42	49	25	15	11	0
CCEI	33	0	67	100	0	0	0	0
CCBA	17	28	65	62	6	10	12	0

* Períodos: 1=agosto-diciembre 2014, 2=enero-mayo 2015

Fuente: Elaboración propia.

4.3. Resultados por facultad: Comparación semestral

Específicamente, en las facultades donde se observó un incremento en el porcentaje de profesores con desempeño excelente fueron las de Química, Odontología y Veterinaria (tabla 7).

Tabla 7. Comparación del porcentaje de profesores de acuerdo a su nivel de desempeño por facultad y período de evaluación

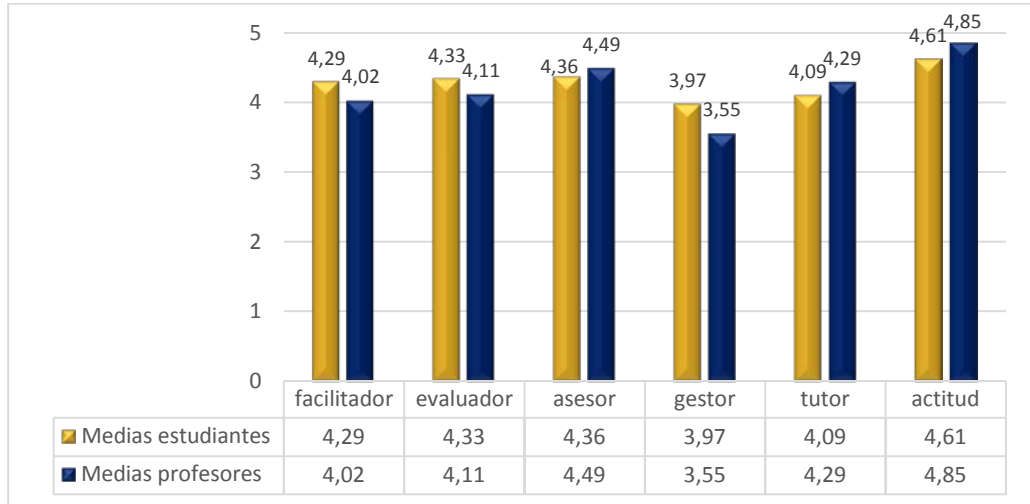
NIVEL DE DESEMPEÑO POR FACULTAD	EXCELENTE		SATISFACTORIO		SUFICIENTE		NECESITA MEJORAR	
	1	2	1	2	1	2	1	2
Período*								
Química	7	13	43	62	29	25	21	0
Odontología	10	58	40	42	10	0	10	0
Enfermería	45	43	45	43	10	14	0	0
Matemáticas	33	0	67	100	0	0	0	0
Veterinaria	17	30	65	60	6	10	12	0

* Períodos: 1 = agosto-diciembre 2014, 2 = enero-mayo 2015

Fuente: Elaboración propia.

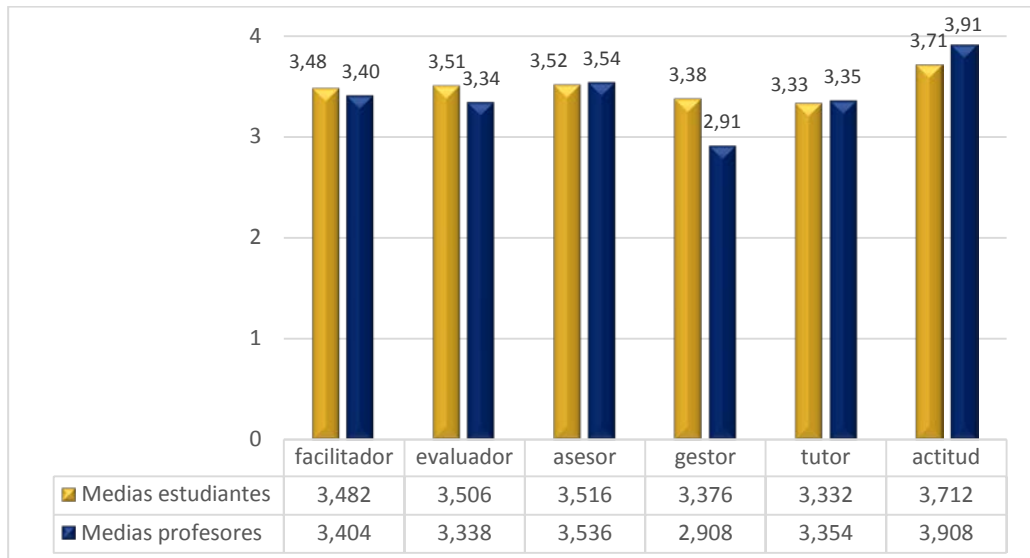
4.4. Comparación de las respuestas de estudiantes y profesores de acuerdo al rol: Período agosto-diciembre 2013

Como se observa en la gráfica 2, existe correspondencia entre la evaluación de estudiantes y profesores. El rol que salió menor puntuado en el período de agosto-diciembre de 2013 fue el de gestor, mientras que el aspecto mejor evaluado es la actitud de los profesores.



Gráfica 2. Comparación de estudiantes y profesores: Período agosto-diciembre 2013
Fuente: Elaboración propia.

Nuevamente, en el período de enero-mayo de 2014 se encontró correspondencia entre las puntuaciones otorgadas por cada rol. En este caso, los profesores se autoevaluaron más bajo en el rol de gestor y tanto estudiantes como profesores evaluaron más alto el aspecto de la actitud.



Gráfica 3. Comparación de estudiantes y profesores: Período enero-mayo 2014
Fuente: Elaboración propia.

4.5. Análisis de comentarios de los estudiantes y profesores

Se categorizaron las respuestas a las preguntas abiertas del instrumento para estudiantes y profesores. Posteriormente, se agruparon de acuerdo a los que tuvieron mayor frecuencia en ambos períodos, agosto-diciembre de 2013 y enero-mayo de 2014.

Respecto a las fortalezas del profesor, las categorías más sobresalientes resultaron las siguientes:

- a) *Interés en el estudiante.* Esta categoría agrupa los comentarios que hacen referencia a la preocupación del profesor porque el estudiante comprenda los temas y actividades. También hace referencia al diálogo que promueve el profesor, así como consejos para mejorar el desempeño en la asignatura. Algunos comentarios representativos son:

Se interesa en los estudiantes y puedes hablar con ella cuando quieras, está interesada en que aprendamos bien.

Es atento con los estudiantes, es buena persona y sabes que puedes hablar con él. Siempre tiene tiempo para nosotros.

Dialoga con el grupo y nos da consejos para un mayor desempeño y aprovechamiento de la materia.

- b) *Dominio del tema.* Esta categoría agrupa los comentarios que hacen referencia a la explicación de los contenidos, la ejemplificación con situaciones reales y la experiencia de los profesores en el tema. Algunos comentarios representativos:

Tiene dominio y habilidad para poder explicar bien el tema.

Presentación clara y desarrollo efectivo de las competencias.

Conoce muy bien sus temas y es muy bueno explicándolos, apoya en horas extras si tienes dudas, era muy atento a todos y es divertido.

Cabe mencionar que esta fue una categoría que se comparte entre los comentarios de estudiantes y profesores; en el caso de los profesores algunos comentarios representativos:

Actualización en la asignatura que imparto y la relación de mis temas con los de otras asignaturas.

Capacidad para relacionar los contenidos con las competencias y el desarrollo de habilidades para la vida profesional.

Experiencia y conocimiento para crear escenarios reales de aprendizaje.

- c) *Habilidad para resolver dudas y problemas.* Esta categoría agrupa los comentarios hechos por los estudiantes donde hacen referencia a la habilidad de los profesores para resolver dudas, tanto dentro como fuera del aula y por vías presenciales y no presenciales. Algunos comentarios representativos:

Resolvió dudas, lo que no entendías te lo volvía a explicar.

Siempre buscaba la manera de que entendamos mejor las cosas, dando asesorías con ayuda de otros profesores, en las salas de cómputo.

Siempre está disponible para ayudarnos a aclarar dudas dándonos asesoría incluso en horas que no son las de la clase.

- d) *Actitudes y habilidades.* Esta categoría agrupa los comentarios que hacen referencia a las conductas, actitudes y habilidades que los estudiantes perciben como son la motivación, la puntualidad, la equidad en el trato y el manejo del grupo. Algunos comentarios representativos:

Su paciencia y buen carácter nos fueron de mucha ayuda para poder avanzar en el estudio y comprensión en la materia.

Evalúa de manera justa los trabajos, en general es muy buen docente.

Es un profesor muy organizado, tiene la plataforma en buen estado.

Es un maestro atento y accesible, es excelente persona.

Demuestra gusto por la materia, se veía en que sus clases eran muy interesantes, eran didácticas y resolvía tus dudas.

Tiene carisma y es comprensivo.

En el caso de esta categoría, existe correspondencia en la opinión proporcionada por los profesores, algunos de los comentarios son:

Paciencia, tolerancia, disposición al diálogo y flexibilidad.

Motivar a los estudiantes para incorporarse a actividades fuera de la universidad para complementar su formación.

Empatía hacia los estudiantes, interés porque aprendan y actitud de respeto.

- e) *Estrategias y actividades de enseñanza y aprendizaje.* Esta categoría agrupa los comentarios que hacen referencia a los aspectos positivos en cuanto al manejo de estrategias de enseñanza y aprendizaje como son el uso de actividades interesantes y significativas, incluir a la planeación didáctica más prácticas y salidas de campo, así como considerar actividades tanto individuales como grupales. Cabe señalar que esta categoría fue más representativa para profesores que para estudiantes. Algunos comentarios representativos de los profesores son:

Las actividades de aprendizaje variadas a las que se enfrentaron los estudiantes como prácticas de laboratorio y de campo.

El empleo de estrategias de enseñanza y aprendizaje significativos.

Gestionar actividades de aprendizaje en la asignatura y garantizar que las clases se cumplieron en tiempo y forma.

Sobre las áreas de mejora, se generaron las siguientes categorías:

- a) *Explicación de los temas y resolución de dudas.* Agrupa los comentarios hechos por los estudiantes donde hacen referencia a los aspectos que puede mejorar el profesor en sus explicaciones. Algunos de los aspectos mencionados son el uso de términos no conocidos, ampliar las aportaciones que los estudiantes realizan en las exposiciones y desarrollar ejemplos durante las mismas. Algunas sugerencias son:

Que el lenguaje que utiliza en cada clase sea un poco más fácil de entender ya que en ocasiones se le olvida que somos iniciados en la carrera.

Que explique el tema directo y no le dé muchas vueltas y se vaya por otro tema.

Que no sólo se pare a hablar y explicar de un Word con el proyector, sino que al mismo tiempo que explique resuelva el ejercicio con una voz más potente.

Explicar más los procedimientos para realizar los problemas matemáticos, los ejercicios y tareas a realizar.

- b) *Estrategias y actividades de enseñanza y aprendizaje.* Esta categoría agrupa los comentarios que hacen referencia a los aspectos que el profesor puede mejorar en cuanto al manejo de estrategias de enseñanza y aprendizaje como son el uso de actividades más dinámicas, proyección de videos, incluir a la planeación didáctica

más prácticas y salidas de campo, así como considerar actividades tanto individuales como grupales. Algunos comentarios representativos son:

No dejar tantas ADAS cuando no hay sesiones presenciales.

Que relacione los temas vistos en clase con nuestro futuro medio laboral.

Que use videos, prácticas en el laboratorio y ejemplos en la vida cotidiana.

Que busque nuevas formas para aplicar los temas, no solo exponer el tema.

Una mejor distribución de ADAS durante todo el semestre.

Esta es una categoría compartida por estudiantes y profesores, al respecto los profesores comentaron:

Generé muchas actividades de aprendizaje.

Son demasiadas materias y demasiados trabajos que les marcamos los maestros.

Requiero capacitación en la búsqueda de nuevas estrategias de enseñanza.

- c) *Estrategias de evaluación.* Esta categoría agrupa los comentarios que hacen referencia a los aspectos que puede mejorar en cuanto a la evaluación, como son aclarar el valor, criterios de evaluación y fechas de entrega de cada actividad, así como brindar retroalimentación de las mismas, no solo al final del semestre. Algunos comentarios son:

Aclarar el valor de las actividades y fechas para mantenernos al tanto.

Que dé retroalimentación de las actividades mucho antes, no hasta que acabe el semestre cuando ya registró la calificación.

Que se evalúe el conocimiento de los alumnos periódicamente para que el profesor esté atento y tenga conocimiento en lo que falta explicar.

Que cada vez que entre al salón, entre haciendo preguntas de lo que vieron en la clase anterior y con eso evaluar quién leyó.

Nuevamente, esta es una categoría compartida por la opinión de profesores y estudiantes. Al respecto los profesores comentaron:

Para mejorar las condiciones de enseñanza es necesario capacitarme en el diseño de instrumentos de evaluación.

- d) *Actitudes y habilidades.* Esta categoría agrupa los comentarios relacionados con algunas acciones que se pueden mejorar como son el uso del celular en el salón de clase, la paciencia para responder y atender dudas, la atención a todos los estudiantes y no solo a un grupo particular, resaltar el valor y utilidad de cada asignatura, el manejo del grupo y cuidar la puntualidad y la asistencia a las sesiones presenciales. Algunos comentarios:

Que no sea tan dura, necesitamos más apoyo que llamadas de atención.

Que sea un poco más estricto.

Exigirles más a los alumnos y menos a él mismo, entiendo que él sabe mucho, pero debe dejar que hagamos algo o no aprenderemos.

Sugiero que debe tener más carácter al dirigirse a los alumnos, ya que en su mayoría, fueron muy groseros con los demás estudiantes en el aula y con él, pues no prestaban atención y hacían lo que querían, perjudicando a los que tratábamos de poner atención a la asignatura.

- e) *Manejo del tiempo.* Esta categoría agrupa los comentarios hechos por los estudiantes donde hacen referencia a los aspectos que podrían considerarse en

las planeaciones respecto a la distribución de los tiempos para el desarrollo de los temas y la resolución de actividades de aprendizaje.

La materia juntaba muchos conocimientos para un semestre. Debería haber otros semestres en los cuales se dé para que en verdad podamos aprender.

Esta asignatura requería de mayor tiempo en actividades presenciales, para despejar interrogantes y la participación de los estudiantes.

El tiempo que se otorga a esta asignatura es poco.

- f) *Uso de la plataforma.* Esta categoría agrupa los comentarios que hacen referencia a los aspectos que podrían considerarse respecto al uso de la plataforma como un medio de interacción entre profesor y estudiantes, así mismo que se use como una herramienta didáctica para variar el estímulo y dinámica de las clases. Algunos comentarios:

Que mejore en el uso de la plataforma y la asignación de calificaciones.

El uso de la plataforma para poner recursos materiales que nos ayuden con los temas y no solo se use para subir tareas.

Que use un poco más la plataforma como apoyo para las actividades, utilizar más TICS en el planteamiento de los temas ya que la dinámica de la clase resultaba muy monótona porque se plantea con la pizarra únicamente.

Emplear y promover las TIC, mayor uso de la plataforma UADY Virtual.

Esta es una categoría que se comparte desde la perspectiva tanto de profesores como de estudiantes, algunos comentarios obtenidos en la autoevaluación son los siguientes:

Necesito más tiempo para optimizar la estructura y actividades plasmadas en la plataforma.

Para mejorar el aprendizaje se debería contar con equipo actualizado de computación y que la navegación de red sea rápida.

Requiero mayor capacitación en el uso de la plataforma.

4.5.1. Resultados recabados con el portafolio de evidencias del profesor

De acuerdo a las declaraciones realizadas en el apartado de reflexión, se obtuvieron los siguientes resultados:

- a) *Compromisos.* Los profesores se comprometen a:
- ✓ Asistir a cursos de capacitación en el área disciplinar.
 - ✓ Perfeccionar sus actividades de aprendizaje para que sean más integradoras.
 - ✓ Capacitarse en el uso de Tecnologías de la Información y Comunicación, particularmente en el uso de la plataforma UADY VIRTUAL.
 - ✓ Capacitación en la implementación del MEFI.
 - ✓ Actualización en el uso de estrategias de enseñanza y aprendizaje.
 - ✓ Entrega oportuna de la retroalimentación de las actividades de aprendizaje y uso de rúbricas más específicas y acordes con las actividades.
 - ✓ Reflexionar sobre los posibles errores que pudieron haberse cometido durante el curso y replicar las acciones que resultaron positivas para el desarrollo de las competencias planteadas en la asignatura.
 - ✓ Fortalecer los roles establecidos en el MEFI.
 - ✓ Trabajar en equipo.
- b) *Áreas de mejora.* Las áreas de mejora identificadas son:

- ✓ Capacitación en el uso de las TIC.
 - ✓ Capacitación en el uso de las estrategias de enseñanza y aprendizaje.
 - ✓ Reconsiderar la ponderación de las actividades.
 - ✓ Reestructurar las estrategias de evaluación.
 - ✓ Mejorar los instrumentos de evaluación.
 - ✓ Capacitación para desarrollarse mejor de acuerdo a los roles establecidos en el MEFI.
 - ✓ Incluir bibliografía en inglés.
- c) *Factores que dificultan el desempeño del profesor.* La información proporcionada por el profesor se divide en factores internos y externos, tal como se presenta en la tabla 8.

Tabla 8. Factores internos y externos que dificultan el desempeño del profesor

INTERNOS	EXTERNOS
1. Cantidad de estudiantes por grupo.	1. Inestabilidad de la plataforma UADY VIRTUAL.
2. Diversidad de estudiantes y sus conocimientos previos.	2. Fallas de internet e interrupciones en el servicio eléctrico.
3. Interés de los estudiantes.	3. Fallas en el equipo de cómputo.
4. Estrategias de enseñanza y aprendizaje utilizadas.	4. Programas extensos, en algunas asignaturas se unieron los contenidos de dos materias del antiguo plan.
5. Número excesivo de actividades de aprendizaje.	5. Que el bachillerato no contribuye al desarrollo de algunas competencias necesarias para ciertas asignaturas.
6. Elaboración de instrumentos de evaluación adecuados.	6. Dificultades familiares y económicas a las cuales se enfrentan los estudiantes e influyen en su desempeño académico.
	7. Las actividades de habilitación del MEFI interfirieron con el tiempo destinado a su labor docente.

Fuente: Elaboración propia

5. Conclusión y discusión

Dentro de las principales aportaciones que se pudieron encontrar al presentar la descripción de la implementación de la evaluación, es favorable observar que incrementó el porcentaje de profesores evaluados con nivel satisfactorio y excelente, elevando de un 74% a un 89%. Otro aspecto favorable es que, para el segundo período en que se realizó la evaluación, no hubo profesores ubicados en el nivel necesita mejorar. Esto puede reflejar la utilidad de la evaluación para retroalimentar el trabajo del profesor, ya que le permitió ubicar sus fortalezas y áreas de mejora y hacer los ajustes necesarios en el segundo semestre, en lugar de ser tomado con un mero elemento para tomar decisiones, en ocasiones financieras exclusivamente.

Otro aspecto que es importante señalar es la correspondencia que se observa en las respuestas otorgadas tanto por estudiantes como por profesores al evaluar el desempeño en cada uno de los roles. Esto indica en primera instancia que la evaluación se está realizando conscientemente y se analizan por ambos actores. El rol que tanto profesores como estudiantes evaluaron más bajo en ambos períodos fue el de gestor, por tanto, surge la necesidad de acompañar a los profesores para esclarecer y proponer actividades que les permitan cumplir con este rol.

Por otra parte, un aspecto que tanto profesores como estudiantes evaluaron más alto en ambos períodos es la actitud, lo que puede indicar la buena disposición de los profesores para generar espacios de respeto, diálogo y promoción de los valores universitarios.

Esto va a la par de los comentarios realizados por los estudiantes y profesores cuando mencionan que algunas de las fortalezas son sus actitudes positivas, así como el interés en el estudiante; es decir, los estudiantes perciben a sus profesores como personas preocupadas por su aprendizaje, a los cuales pueden acudir y estos los tratarán paciente y respetuosamente. Por su parte, los profesores coinciden con esta perspectiva al considerarse profesionales que permiten el diálogo, se interesan por sus estudiantes y generan un ambiente de respeto. Estos elementos a su vez se relacionan con el rol de tutor del profesor de acuerdo a lo establecido en el MEFI (UADY, 2012).

Otro punto en el que coinciden estudiantes y profesores es en la percepción de estos como expertos en los contenidos de sus asignaturas, en su preparación académica y, además, por su experiencia docente. Esto es acorde a lo estipulado en el Plan de Desarrollo Institucional (PDI, 2010), en donde se establece que todos los académicos de la universidad se caracterizan por su amplia experiencia y preparación profesional, lo que les permite desarrollar competencias en la impartición del Modelo Educativo de la Universidad.

Sin embargo, también se identificaron áreas de mejora en la calidad de las explicaciones y en un aspecto en el que coinciden estudiantes y profesores es la necesidad de variar el uso de estrategias y actividades para hacer más dinámico el proceso de enseñanza y aprendizaje. Estos aspectos se relacionan con el rol de facilitador y a partir de estos resultados se está generando un programa de actualización y capacitación que brinde herramientas a los profesores.

De nueva cuenta, profesores y estudiantes consideran que los docentes deben ser más precisos en el uso de sus estrategias de evaluación, definir claramente sus criterios y proporcionar retroalimentación. A su vez, se identificaron otras áreas de mejora como actitudes poco favorables, como la falta de paciencia, el trato equitativo hacia los estudiantes, mejorar las habilidades en el manejo del grupo, así como mejorar la puntualidad y la asistencia a las sesiones presenciales. La atención de estas áreas de mejora puede considerarse una oportunidad en el proceso de transformación institucional y en el marco de la responsabilidad social universitaria al favorecer la mejora del perfil del profesorado, fomentando la conformación de una planta académica con altos niveles de habilitación, de acuerdo a lo establecido en el Plan de Desarrollo Institucional 2014-2022 (PDI, 2014).

De esta información, se generan recomendaciones en general y en específico que se trabajan al interior de las facultades. Algunas de estas son: manejar mejor los tiempos para dar oportunidad de abarcar todos los temas establecidos en la planeación didáctica, integrar a esta actividades de aprendizaje diversas e inclusive utilizar la plataforma virtual como recurso de apoyo en el proceso de enseñanza y aprendizaje. Otro aspecto importante a considerar es brindar retroalimentación periódica de las actividades y evaluaciones realizadas con el fin de monitorear sus avances durante el semestre.

Es interesante apuntar que las áreas de mejora identificadas por los estudiantes y profesores también coinciden con las áreas mencionadas por los propios profesores en su portafolio, ya que estos solicitan capacitación en el uso de las tecnologías de información y comunicación (TIC), en especial de la plataforma virtual, así como en el diseño e implementación de estrategias de enseñanza-aprendizaje y evaluación. De modo que los profesores se comprometen a capacitarse en aspectos disciplinares, lo concerniente al Modelo Educativo, en el uso de las TIC. De igual forma, a continuar actualizándose el ámbito pedagógico para mejorar el diseño de sus planeaciones

didácticas, el diseño de actividades de aprendizaje más integradoras y mejorar los instrumentos de evaluación utilizados en el proceso de enseñanza y aprendizaje.

Por último, es necesario reflexionar sobre los factores internos y externos que dificultan el desempeño del profesor. En el caso de los factores internos, se encuentran en primer plano el número de estudiantes, así como los conocimientos previos, características e intereses de estos. Dada la naturaleza pública de la universidad, sería poco viable la reducción de estudiantes aceptados, sin embargo, es posible la capacitación de los profesores para el manejo de grupos numerosos, la generación de estrategias acordes a las características de los grupos, así como la reducción de las actividades de aprendizaje. Es decir, cambiar la perspectiva de las “tareas” al diseño de experiencias de aprendizaje más integradoras con criterios de evaluación claramente establecidos, esto con el fin de facilitar el trabajo de los profesores en el curso de sus asignaturas.

Por otra parte, se encuentran los factores externos como son la inestabilidad de la red, lo que dificulta el uso de la plataforma virtual, aunado a las fallas del equipo de cómputo de las facultades. Además, al realizarse la reestructura de los planes de estudio, algunas asignaturas se fusionaron en una sola, por lo que los profesores perciben que los programas son más extensos y no cuentan con el tiempo necesario para abarcar los contenidos. Es importante trabajar en la capacitación de los profesores relativa al desarrollo de competencias, ya que continúan dando mayor importancia al contenido, el cual es solo un elemento de la competencia y no dimensiona el trabajo integral que implica el desarrollo de competencias.

Referencias

- Aguilar, V. (2009, septiembre). Sistema de evaluación docente. Hacia un modelo formativo e integral en educación superior. Ponencia presentada en el *X Congreso Nacional de Investigación Educativa*, Veracruz.
- Bonifaz, R. (2011). Origen de la evaluación docente y su conexión con las políticas públicas en educación. En J. Manzi, R. González y Y. Sun (Eds.), *La evaluación docente en Chile* (pp. 13-34). Ciudad de México: Centro de Medición MIDE.
- Cisneros, E. y Aguilar, Z. (2010). Validación de un modelo de desarrollo y evaluación de competencias académicas en el sureste de México. En E. Cisneros, B. García, E. Luna y R. Marín (Coords.), *Evaluación de competencias docentes en Educación Superior* (pp. 243-266). Ciudad de México: REDECA.
- De Miguel, M. (1998). La evaluación del profesorado universitario. Criterios y propuestas para mejorar la función docente. *Revista de Educación*, 315, 67-83.
- García, B., Loredó, J., Luna, E. y Rueda, M. (2008). Modelo de Evaluación de Competencias Docentes para la Educación Media y Superior. *Revista Iberoamericana de Evaluación Educativa*, 1(3e), 124-136.
- Rizo, H. (1999). Evaluación del docente universitario. *Revista Electrónica Interuniversitaria del Formación del Profesorado*, 2(1), 425-440.
- Rueda, M. (2008). La evaluación del desempeño docente en la universidad. *Revista Electrónica de Investigación Educativa*, 10, 1-15.
- Hesse-Biber, S. (2010). *Mixed methods research: Merging theory with practice*. Nueva York: Guilford Press.

SEP (2010). *Lineamientos de evaluación docente. Evaluación del desempeño docente bajo el enfoque por competencias*. Ciudad de México: Dirección General de Bachillerato.

UADY. (2012). *Modelo Educativo para la Formación Integral*. Mérida (Yucatán): Autor.