

Dynamic Assessment of Reading Difficulties

Evaluación dinámica de las dificultades de aprendizaje de la lectura

Juan-José Navarro and Joaquín Mora

University of Seville¹

Abstract

This article describes the changes produced in a group of students with reading difficulties (RD) in response to the experimental application of a dynamic assessment device of the processes involved in reading (EDPL). The sample consists of 133 students of 9-16 years with RD (60 in the experimental group and 73 in the control group). The design is a pretest/posttest for impact evaluation. Analyses allow differentiated information on the impact of the EDPL device depending on the previously established degree of difficulty; empirically setting the population that would get more benefits from its application in the school context. The results show significant improvements in reading comprehension, personal-social adjustment and intelligence in the students who have greater difficulties both in specific domains and in intelligence. Likewise, the analysis provides information about the mechanisms that might explain the improvements.

Keywords: Dynamic assessment, reading comprehension, reading difficulties.

Resumen

En este artículo se recogen los cambios producidos en un grupo de estudiantes con dificultades de comprensión lectora (DL) como respuesta a la aplicación experimental de un dispositivo de evaluación dinámica de procesos implicados en la lectura (EDPL). La muestra está integrada por 133 estudiantes de 9 a 16 años con DL (60 en el grupo experimental y 73 en el grupo control). Se utiliza un diseño pretest/postest para la evaluación de los resultados. Los análisis efectuados permiten obtener información diferenciada sobre el impacto de EDPL en función del grado de dificultad previamente establecido, configurando empíricamente la población que obtendría mayores beneficios de su aplicación en el contexto escolar. Los resultados muestran mejoras significativas en comprensión, ajuste personal-social e inteligencia en aquellos estudiantes que presentan mayores dificultades, tanto en dominios específicos como en inteligencia. Así mismo, los análisis otorgan información acerca de los mecanismos que podrían explicar las mejoras.

Palabras clave: Evaluación dinámica, comprensión lectora, dificultades de aprendizaje de la lectura.

Correspondence: Juan José Navarro Hidalgo, Departamento de Psicología Evolutiva y de la Educación, Universidad de Sevilla, c/ Pirotecnia, s/n., 41013, Sevilla (Spain). E-mail: jnavarro@us.es.

¹ We would like to express our gratitude to the academic staff at the different educational centres participating in this research, as well as to those at the Teachers' Centre (CEP) in Alcalá de Guadaíra (Seville), for their valuable collaboration. Likewise, we thank the Department of Education of the Andalusian Government the granting of a study leave for this research.

Introduction

Dynamic assessment models are based on the interaction between the subjects being assessed and other more expert subjects, who provide help, establishing a mediation process between current competences and task demands. Thus, a space in which to build knowledge is created, which is particularly sensitive to pedagogic action, hereby facilitating the identification of the type and degree of educational aid needed, and establishing a dynamic measure of the processes assessed, in contrast to more static measures that basically inform about skills that are already consolidated (Elliott, 2003; Grigorenko, 2009; Newman, Griffin, & Cole, 1989; Sternberg & Grigorenko, 2002; Vygotsky, 1934/1995). The application of these assessment models seems to be more relevant with students who present learning difficulties (LD) or cognitive disability (Campioni & Brown, 1987; Carlson & Wiedl, 2000; Guthke & Beckmann, 2000; Saldaña, Mora, & Moreno, 2007; Swanson & Howard, 2005). However, the consideration of the social context characterized by the interaction and the support in which the educational process normally evolves can contribute significantly to the increase of the ecological validity of these procedures, facilitating their application and generalization in the school setting (Schlee, 1985; Thurman & McGrath, 2008; Wells, 1999).

One of the main reasons branched in favor of dynamic assessment (DA) is its capacity to identify interindividual differences while minimizing the influence of students' previous history (Kozulin, 1998; Resing, Tunteler, de Jong, & Bosma, 2009). In effect, as it consists of activities that attempt to assess the process and capacity to learn certain concepts or skills, establishing a common starting point concerning the contents that are dealt with and focusing especially on skills and competences *under development*, these procedures would no doubt be rated better than conventional tests (Dörfler, Golke, & Artelt, 2009; Lidz & Elliot, 2000; Sternberg & Grigorenko, 2002). Some studies have shown that children with very different schooling histories could perform equally well in tests that assessed their learning potential (Feuerstein, 1996; Kan-iel, Tzuriel, Feuerstein, Ben-Shachar, & Eitan, 1991). There is also evidence that students with special difficulties improved their performance significantly in assisted or collaborative learning situations, as well as after the application of dynamic assessment procedures, and these works generally show positive effects in the treatment groups (Campioni, Brown, & Ferrara, 1987; Fernández-Ballesteros & Calero, 2000; Guthke & Beckmann, 2000; Resing, 2000; Resing & Elliott, 2011; Resing et al., 2009; Saldaña et al., 2007). Thus, significant improvements in specific do-

mains, such as working memory, arithmetic or reading, have been reported (Swanson & Howard, 2005), as well as in general skills measured with IQ tests after applications focused essentially on metacognitive processes, in which the application conditions of the tasks were manipulated with regard to the quantity or type of verbalization and feedback received by the subject during the mediation (Beckmann, Beckmann, & Elliott, 2009; Carlson & Wiedl, 2000).

However, DA allows access to the integrated assessment of diverse processes involved in an activity (Rapp & Broek, 2005). Thus, the information about the mechanisms through which a student builds and optimizes his learning can be relevant to orient the educational action (Ruiz & Camps, 2009), thereby allowing the establishment of real connections between the assessment and intervention processes, considered a key aspect to improve educational processes (Aguilera & García, 2004; Coll, Barberá & Onrubia, 2000; Gersten & Dimino, 2006; Grigorenko, 2009). In his proposals about assessment, Vygotsky made a particular distinction between DA in general cognitive functions (the more developed sphere) and DA in specific academic learning domains, such as reading or arithmetic (Kozulin & Garb, 2001). In our opinion, the nature of the tasks proposed from each one of these approaches is particularly relevant for education. Thus, on the one hand are the

dynamic tests that include *new tasks for the student*; these tasks are characterized by their decontextualization and their similarity to conventional tests; they attempt to assess learning capacity (Kozulin, 1998; Sternberg & Grigorenko, 2002). Among their main advantages are: (a) they assess the learning potential in “culture-free” tasks and (b) the supposed equality of conditions when facing the tasks. Among their drawbacks are: (a) they assess “abstract” learning potential, with no correspondence with learnings that are meaningful for school, (b) the decontextualization of the activity of learning, (c) the scarce/lack of ecological validity, and (d) the supposedly lower predictive power of performance as a consequence of these elements. On the other hand are the tests that include *tasks with contents related to academic learnings* (Cioffi & Carney, 1983; Dörfler et al., 2009; Grigorenko, Sternberg, & Ehrman, 2000; Haywood & Lidz, 2007; Sternberg & Grigorenko, 2002). They have the following advantages: (a) the contextualization of the assessment instruments, (b) greater ecological validity of the results, and (c) the supposedly higher prediction power of academic performance. They also have the following drawbacks: (a) as they are activities whose contents are saturated with information that may already be a part of the subject’s knowledge, it would be more difficult to establish the learning capacity as such and, as a

consequence of this, (b) the need to exhaustively establish the initial level of development to be able to determine the improvements, as well as the need to build specific assessment devices, sensitive to the processes involved, and efficacious in the school setting (Lidz, 2004; Thurman & McGrath, 2008).

Goals and research hypotheses

The present work attempts to analyze the effects in a group of students with reading comprehension difficulties (RD) by the experimental application of a dynamic assessment device of the processes involved in reading (in Spanish, “evaluación dinámica de procesos implicados en la lectura” [EDPL]). Our hypothesis contemplated possible improvements in comprehension, as well as in the levels of personal-social adjustment and cognitive performance. This hypothesis is based on the processes of *mediated interaction* that take place during the application, as well as on the metacognitive content that underlies most of the mediation guidelines provided by the test administrators during the sessions.

Method

Participants

Thirteen schools (7 from primary education and 6 from compulsory secondary education [CSE]),

located in Seville and Cadiz (Andalusia-Spain), participated the investigation. The sample comprised 133 students with RD, of whom 60 (23 girls and 37 boys) were in the experimental group (EG) and 73 (29 girls and 44 boys) were in the control group (CG). Age range was between 9-16 years. The students were assigned to these groups as a function of the participation possibilities of the teachers who collaborated with the Work Group (WG) that implemented the EDPL device. There were 12 teachers/counselors (test administrators) in this WG, and another 5 were in the CG. The lower number of members in the CG is due to the fact that some counselor-participants of the CG attended large groups of students and, likewise, some counselors from the EG carried out their work in centers with students who ultimately made up the CG. In both cases, they were professionals involved in innovation projects, with training and experience in special education, psychology, or pedagogy. Forty-eight teacher-tutors collaborated in completing the rating scales, and 6 teachers who did not belong to the centers participated in the application of the criterion-tasks. Initial sample selection was carried out on the basis of the information about the significant presence of RD provided by the specialist teachers and counselors in a data-collection form. Various subgroups were contemplated as a function of the information provided by the

educational counseling teams and by the centers, which were represented in each one of the experimental conditions. These subgroups were not exclusive, and their con-

sideration in the subsequent analyses should contribute to better delimiting the efficacy sphere of the *EDPL* device. The details of the sample are in Table 1.

Table 1

Composition and Distribution of the Sample

Distribution by stage and participation groups				Participant teachers		Subgroups contemplated for analysis	EG	CG
Stage	Students	EG	CG	EG (members of the WG)	12	Learning difficulties	31	47
Primary	68	32	36	CG	5	Cognitive disability	13	16
	65	28	37	Experts in validation	4	Socioculturally deprived	16	10
CSE				Application of criterion-tests	6	<i>Receives specific support</i>	39	26
				Collaborating teachers	48			
Totals	133	60	73					

Note: The subgroup *Receives specific support* was made up of students with RD who received specialized educational attention, and it included subjects from the other subgroups.

Instruments

The EDPL device

EDPL (Navarro, 2008) is a set of strategies and activities of DA through which information about the process of resolving the proposed tasks is obtained, as well as of the mechanisms that contribute to improvement during the activity. In accordance with the theoretical framework of reference adopted, the proposed reading assessment is basically characterized by considering reading an interactive (Graesser, Singer, & Trabasso, 1994; Rapp &

Broek, 2005), systemic and dynamic process, which is socially built in interaction contexts (Wells, 1999), based on a process of *mediated interaction* between the assessor and the student with difficulties, thereby providing a connection between assessment and intervention (Aguilera & García, 2004; Grigorenko, 2009). In this sense, the assessors take on an active role in the assessment situations, they modify their questions and even the instruments as a function of the subject's response; they provide tools or guidelines to control and regulate the activity in order to observe and

appraise their degree of incorporation into the task resolution; they intervene in the process, mediating its accomplishment, in order to assess the process of performance and the degree and type of support the subject needs (Annex I).

The *EDPL* device is made up of 32 assessment activities (Fig. 1) placed within a series of processes that have been contemplated as a

function of prior research (Compton et al., 2010; Graesser et al., 1994; Hacker, 1998; Kintsch, 1988; Kintsch & Kintsch, 2005; Sánchez, 1999). Each activity includes: (a) the process to be assessed; (b) a description of the activity to be carried out; (c) the methodology proposed for its application; (d) the patterns of mediation and assessment of metacognitive processes, that is, the

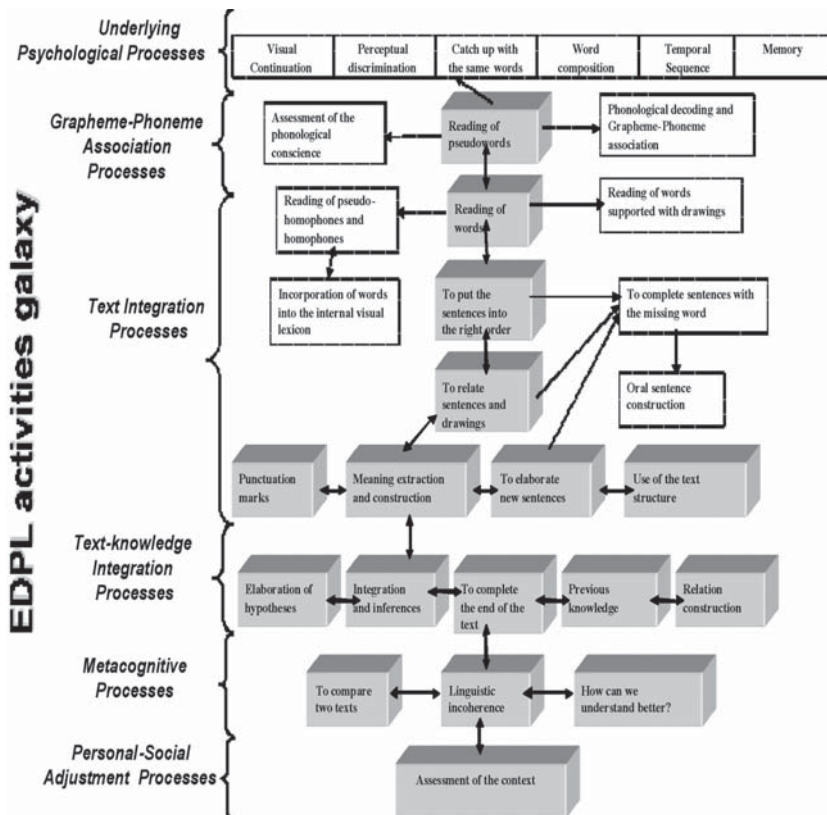


Figure 1. Assessment activities of the diverse reading processes contemplated in EDPL (Reprinted from Navarro, & Mora, 2011).

tools and guidelines to control and regulate the aforementioned activity; and (e) assessment criteria or indicators.

To facilitate collecting and appraising the information, the test administrators had *Recording and assessment sheets* for each activity. They should reflect the process of mediation and assessment carried out on these sheets. This assessment consisted of the analysis of the process of learning and resolving tasks, which involved difficulties about the “invisibility” of the cognitive and emotional processes at stake, the deployment of which the evaluator necessarily had to infer. A list of indicators (Table 2) was elaborated to facilitate the task of assessing more accurately the presence and quality of the diverse

processes. These indicators were handed in to the test administrators as a part of each one of the activities and should be appraised on the Recording and assessment sheets. A system of analysis was developed to quantify the assessments carried out (Lukas & Santiago, 2004). This allowed us to operationalize the process of application and to obtain *dynamic scores* for each activity and also global scores. The system adopted (Moreno & Saldaña, 2005), was based on a gradual scale made up of four appraisals: (1) The behavior does not appear. The indicator was not deployed; (2) Signs, rudimentary behavior; (3) Clearly observable behavior (although not high quality); and (4) Intense behavior or high-quality behavior.

Table 2

Indicators of the Block of Processes of Text-Prior Knowledge Integration

He/she integrates new ideas with prior knowledge and experiences

He/she uses prior knowledge and experiences to construct meanings

He/she is sensitive to the author’s style, assumptions, perspective and intentions and can even question these aspects

He/she expresses opinions, judgments, or intuitions about the text contents

He/she establishes relations among diverse texts

He/she formulates and answers questions about the text

He/she formulates hypotheses. He/she anticipates results

He/she contrasts and verifies the hypotheses

He/she makes inferences about the text that go beyond it

He/she reviews the ideas of the text based on what he/she already knows

Criterion battery

- *Factor “G” test of Cattell, Scale 2-Form A* (Cattell & Cattell, 1974/2001).
- *Scale of Personal-social adjustment regarding Reading* [in Spanish, Escala de Ajuste Personal-Social con relación a la Lectura, *APSLJ*]. This scale (Navarro, 2008) provides information about 9 dimensions (Table 3) and is made up of 80 items presented within a structure of short sentences on which the students should rate their degree of agreement or disagreement. A global score and scores of the diverse dimensions are obtained.
- *Text comprehension*: We used two tests to assess text comprehension: *ECO1*, for the second and third cycles of primary education, and *ECOS*, to be applied in CSE (Table 4). They were elaborated in view of the goals of the study (Navarro, 2008) and they assess the student’s knowledge and use of reading strategies, as well as global comprehension. The *ECO1* and *ECOS* comprise 10 texts with diverse structures, belonging to various spheres related to the curricular areas of primary and secondary education, respectively, followed by a set of multiple-choice questions (43 for *ECO1* and 56 for

Table 3

Dimensions of the APSL Scale

-
- D1 Knowledge of one’s own current cognitive limitations, potential, and possibility of learning, difficulties, and ways to solve them.
 - D2 Knowledge of one’s expectations, academic self-concept, and attributional style.
 - D3 Knowledge of one’s attitude toward reading: emotions, feelings of rejection *versus* satisfaction and enjoyment of reading, frustration tolerance, persistence in the face of difficulty, autonomy *versus* dependence.
 - D4 Knowledge of one’s own interests and motivations.
 - D5 Perception of support provided by the school and family settings, and of the conditions in which this occurs.
 - D6 Perception of the relevance of the topic: reading, what it is good for, its importance.
 - D7 Knowledge of one’s own relational strategies with others in reading tasks.
 - D8 Knowledge of one’s own strategies of control and conscious regulation of the reading process (planning, supervision, self-assessment).
 - D9 Awareness of the capacity to generalize this knowledge or strategies.
-

Table 4

Dimensions of the ECO1 and ECOS Tests of Text Comprehension

Conception of reading: What is reading good for, what does reading consist of, what does reading imply.

Knowledge of the strategies that can be applied: self-regulation strategies (planning, supervision, self-assessment), transfer and generalization, textual integration, and integration of text-prior knowledge.

Prior knowledge (only in ECOS): Knowledge about the text topic, situations posed, and action outlines described.

Textual integration: access to the meaning of words, linear integration of the ideas, integration of all the ideas in an outline, construction of global propositions, distinction of the relevant information, ranking and sequencing the information, access to the information contained in charts, tables, and outlines, identification and use of the textual structure.

Integration of the text with prior knowledge and experience: establishment of relations, formulation of hypotheses, inferences, author's intention, critical review of the text from what is already known.

Supervision of the process of comprehension: detection of incoherencies and overcoming the difficulties by applying strategies of lexical, syntactic, and semantic processing.

ECOS). The application of these tests provides a global score in comprehension, and scores in the diverse dimensions.

The analyses of the validity and reliability of the comprehension and personal-social adjustment tests, as well as its validation by experts before its application, yielded positive results. These tests were administered to groups of complete classes from the collaborating centers, obtaining, among others, the following data: *ECO1* test ($n = 179$): alpha reliability coefficient = .81, Kolmogorov-Smirnov normality test = .088 ($p < .002$), mean homogeneity index (MHI) = .32, mean difficulty index (MDI) = .65; *ECOS*

test ($n = 43$): alpha = .89, Shapiro-Wilk normality test ($n < 50$) = .94 ($p < .027$), MHI = .37, MDI = .51; *APSL* scale ($n = 98$): alpha = .88, Kolmogorov-Smirnov = .077 ($p < .179$), MHI = .32, MDI = .31. As the *APSL* data were nonnormal, nonparametric tests were used in some of the analyses carried out with small samples.

Design and procedure

A pretest-posttest design with a control group was used in order to assess the impact of the intervention. The procedure followed is summarized in Table 5. The *EDPL* device was applied in a small group. The 12 test administrators

Table 5

Procedure Followed in the Investigation

-
1. Pilot application of the instruments designed and validation by experts.
 2. We contacted the collaborators and families in the experimental phase.
 3. Programming and starting the sessions of the Work Group (WG). Training the test administrators.
 4. Training the collaborators who administered the criterion-tests.
 5. The criterion-tests were applied to the EG and the CG.
 6. The device was applied in the EG by the teachers/administrators /The CG receives *other treatments*.
 7. Follow-up and assessment of the process of application of the *EDPL* by the WG.
 8. Posttest application of the criterion-tests.
 9. Assessment sessions in the WG of the *EDPL* device and its application.
 10. Final data collection. Determination of results as a function of hypotheses.
-

were trained in the theoretical/methodological bases of the proposal. These training sessions took place in a WG coordinated by the main investigator, and the follow-up and assessment of the entire process was carried out through this WG. The specific contents of these sessions included the presentation, discussion and conjoint analysis of the activities proposed, examples of the application of the mediation guidelines presented in the device, analysis of videos, and follow-up of the initial sessions. The application duration was 16 weeks; each application group (a total of 11) invested an average of two weekly 45-minute sessions. According to the proposed goals, the administrators had to perform the activities presented in the

tridimensional rectangles of Fig. 1 in a predetermined sequence. As the CG was made up of students with RD who received academic support, this group was actually a group of *other treatments*, the effect of which would, in any event, be randomized depending on the diversity of actions, centers, and participant teachers.

Table 6 shows the values of initial equivalence of the EG and the CG for diverse variables, including the presence in both groups of subjects from the different subgroups. Generally, no significant group differences were found in the analyses. Likewise, the values of Levene's test of variance homogeneity also showed this initial equivalence. In the contrasts of intelligence, adjust-

Table 6

Analysis of the Initial Global Equivalence of CG and EG by Subgroup for the Variables Intelligence, Adjustment, and Comprehension

Pretest measures	Global		Subgroups			
			LD	DIS	DEP	SS
Intelligence	M-CG (SD)	80.91 (18.05)	87.71 (16.36)	68.25 (14.30)	88.80 (16.83)	70.70 (14.61)
	M-EG (SD)	82.57 (14.00)	83.06 (15.03)	76.31 (10.76)	86.70 (13.20)	79.82 (13.37)
	Levene	3.78	.45	.53	.59	0.18
	Sig.	.054	.50	.47	.45	0.67
	<i>t</i>	(2, 103) = .53	(2, 53) = -1.09	(2, 27) = 1.68	(2, 19) = -.29	(2, 57) = 2.40
	Sig.	.60	.28	.11	.77	.02*
Personal-social adjustment	M-CG (SD)	95.06 (15.90)	95.50 (15.29)	97.38 (11.47)	89.28 (24.02)	92.65 (13.36)
	M-EG (SD)	97.98 (13.91)	95.78 (14.32)	97.76 (12.40)	102.13 (14.31)	94.77 (13.18)
	Levene	.16	.05	.24	3.44	.30
	Sig.	.69	.82	.63	.08	.59
	<i>t</i>	(2, 115) = .98	(2, 63) = .08	(2, 27) = .09	(2, 21) = 1.62	(2, 62) = .63
	Sig.	.33	.94	.93	.12	.53
Comprehension (primary)	M-CG (SD)	86.16 (8.91)	88.75 (8.55)	80.75 (9.04)	82.66 (8.22)	80.47 (7.72)
	M-EG (SD)	84.18 (8.75)	86.92 (8.57)	84.35 (6.78)	81.95 (9.64)	84.79 (9.53)
	Levene	.25	.31	.15	.26	.13
	Sig.	.62	.58	.71	.61	.72
	<i>t</i>	(2, 66) = -.93	(2, 31) = -.58	(2, 9) = .76	(2, 22) = -.19	(2, 26) = 1.14
	Sig.	.36	.36	.36	.36	.36
Comprehension (secondary)	M-CG (SD)	91.17 (8.02)	92.42 (8.20)	88.56 (7.28)	82.41 (4.09)	88.20 (6.75)
	M-EG (SD)	88.25 (10.24)	90.73 (9.68)	79.52 (6.34)	83.14 (12.30)	83.33 (5.67)
	Levene	1.67	.95	1.07	1.59	1.52
	Sig.	.20	.34	.32	.28	.23
	<i>t</i>	(2, 63) = -1.29	(2, 43) = -.64	(2, 16) = -2.59	(2, 4) = .08	(2, 35) = -2.38
	Sig.	.20	.53	.02*	.94	.02*
Presence of subgroups in EG and CG	CG		47	16	10	26
	EG		31	13	16	39
	χ^2		3.28	.31	1.39	2.60
	Sig.		.070	.58	.24	.11

Note: In the contrasts, equal variances were assumed; χ^2 = Chi square; M-CG/EG (SD) = Mean of control group/experimental group and standard deviation; LD = Learning difficulties (excluding deprivation and disability); DIS = Cognitive disability; DEP = sociocultural deprivation (excluding disability); SS = Receives specific support; * $p < .05$.

ment, and comprehension, *Student's t-test* for independent groups were used. With regard to age, the mean values of the EG (12.40, $SD = 2.44$) and of the CG (12.05, $SD = 2.34$), as well as Levene's test and *Student's t-test* revealed equivalence ($F = .59$, $p < .44$; $t_{(2, 131)} = .83$, $p < .41$). The groups were also equivalent in sex (χ^2 EG/CG girls = .69, $p < .41$; χ^2 EG/CG boys = .61, $p < .44$).

Data analysis and statistical treatment

In addition to the teachers' prior designation as a criterion to establish the RDs, in our analysis, the initial scores in the tests of comprehension, adjustment, and IQ were considered, establishing three analysis criteria as a function of the rating of the difficulties: (a) Initial rating by teachers of students with RD; (b) teachers' rating + initial score in Comprehension [or Adjustment] ($\leq 84/100$); and (c) teacher's rating + initial score in Comprehension [or Adjustment] ($\leq 84/100$) + initial score in IQ ($\leq 84/100$). A student could thus meet one criterion, two, or all three criteria. Thus, it was possible to analyze the results from a complementary perspective as a function of the analysis criterion adopted, also allowing us to obtain differential information about the impact of *EDPL* depending on the student's degree of difficulty. All the scores of this investigation were transformed by applying a *standardized scale* with a mean of 100

and SD of 16 (Botella, León, San Martín, & Barriopedro, 2001). For the contrasts with the *APSL* scale in small subgroups, Wilcoxon's non-parametric tests were used for related samples and Mann-Whitney's *U* for independent samples. The values of the effect size for each contrast were calculated with Pearson's point-biserial correlation coefficient $r_{Y\lambda}$ (Pardo & San Martín, 1998). These values provide information not only about the existence of a significant effect of the treatment applied, but also of its impact on the variables studied. The proposed standardization assumes that values of $r_{Y\lambda}$ over .37 reflect a large effect size and, therefore, a significant impact of the treatment applied.

Results

The analyses for each test were carried out as a function of the three analysis criteria established (Table 7), and taking into account the diverse subgroups contemplated (Table 8). We shall refer to the most relevant effects observed in each test-criterion, including some data from the contrasts carried out of the diverse dimensions.

Firstly, it must be admitted that the device had no significant differential effects on performance in comprehension as a function of the criterion of the teacher's prior's designation, that is, when all the subjects were taken into account. Although the inspection of the results

Table 7

Results in CG and EG for the Improvement Score (posttest-pretest) in the Diverse Tests, as a Function of the Three Analysis Criteria Established

Tests		(a) Prior teacher's designation			(b) Teacher's criterion + initial score in comprehension/ adjustment (< = 84)			(c) Teacher's criterion + initial score comprehension/ adjustment (< = 84) + initial score in Cattell (< = 84)		
		IS	n	r _{Yλ}	IS	n	r _{Yλ}	IS	n	r _{Yλ}
ECO1	CG	+6.04*	36	.39	+13.98**	13	.76	+7.10	6	.51
	EG	+0.60	30	.07	+6.12*	11	.60	+6.73*	7	.73
ECOS	CG	+5.15**	34	.56	+8.06*	7	.75	+12.30*	4	.91
	EG	+7.93**	25	.74	+8.68**	12	.84	+5.79*	7	.78
APSL	CG	-2.22	56	-.15	+4.01	12	.31	+10.31	7	.67
	EG	+0.11	54	.00	+8.36*	11	.75	+7.60*	8	.73
Cattell	CG	+3.07	42	.25	-3.66	15	-.35	-3.00	11	-.39
	EG	+4.79**	57	.41	+7.09**	24	.69	+8.33**	15	.74

Note: In the pre-post contrasts, the t-test for related samples was used; the analysis criteria (b) and (c) for APSL contemplate the initial score obtained in the scale; IS = improvement score; * $p < .05$; ** $p < .01$.

Table 8

Results in the CG and EG for the Improvement Score (posttest-pretest) in the Diverse Tests, as a Function of Diverse Subgroups

Tests		LD			DIS			DEP			SS		
		IS	n	r _{Yλ}	IS	n	r _{Yλ}	IS	n	r _{Yλ}	IS	n	r _{Yλ}
ECO1	CG	+5.30	22	.35	+1.68	4	.20	+9.42	10	.52	+1.12	8	.12
	EG	-0.20	11	-.03	+1.28	7	.20	+0.93	12	.09	+0.23	19	.03
ECOS	CG	+5.24**	24	.53	+4.92*	10	.66	—	0	—	+5.15**	16	.75
	EG	+9.44**	17	.77	+5.54*	6	.80	+2.17	2	.51	+8.25**	17	.74
APSL	CG	-2.83	33	-.17	-2.30	16	-.30	+0.80	7	.04	-0.49	26	-.06
	EG	+2.04	29	.16	-4.89	13	-.48	+0.82	12	.09	+0.72	37	.06
Cattell	CG	+4.72	22	.34	+2.26	15	.26	-1.90	5	-.13	+2.69	23	.26
	EG	+2.77	30	.25	+10.64**	11	.74	+4.56	16	.41	+5.58**	36	.46

Note: In the pre-post contrasts, the t-test for related samples was used. LD = learning difficulties; DIS = disability; DEP = socioculturally deprived; SS = receives specific support; IS = improvement score; * $p < .05$; ** $p < .01$.

shows a positive evolution in the improvement of the EG in *ECOI* as the analysis criterion indicated more difficulties, the effects are overshadowed by the results of the CG. Despite the initial group equivalence, the higher presence in the EG of students who received specific support (SS) could have had significant impact on the low post-test scores of this group. This will be further commented on in the *Discussion*. Moreover, by means of the *t*-test for independent samples, the improvement scores for *ECOI* between EG and CG were contrasted. The mean differences were favorable to the CG, and were significant for the dimension *Textual integration* as a function of the first analysis criterion ($t_{(2,64)} = 2.190, p < .032$) and of the second analysis criterion ($t_{(2,22)} = 2.147, p < .043$).

With regard to the *ECOS* test, significant improvement was observed in both groups as a function of the three analysis criteria. The contrasts of the improvement scores showed significant differences in favor of the EG as a function of the first criterion for the dimension *Supervision and self-assessment of the comprehension process* ($t_{(2,57)} = -2.026, p < .047, r_{Y\lambda} = .26$). The EG improved significantly in the analyses carried out for the diverse subgroups, except for the DEP subgroup, although the mean differences with the CG did not reach statistical significance.

With regard to the *APSL* scale, the results show significant im-

provement in the EG as a function of the second and third analysis criteria. These improvements were produced in the global score in the dimensions related to strategic meta-knowledge of reading (D8) and with the awareness of the capacity to generalize strategies (D9) (*Wilcoxon* $_{(n=11)} = 32.00, p < .048$, as a function of criterion (b) and *Wilcoxon* $_{(n=8)} = 26.50; p < .033$ for criterion (c)). We also carried out contrasts for *APSL* as a function of the stage. These analyses allowed us to verify that dimension D9 and the dimension that concurrently expresses dimensions D8 and D9 (D8 and 9) established significant differences in CSE (D9, $t_{(2,25)} = -2.848, p < .009, r_{Y\lambda} = .49$ and D8 and 9, $t_{(2,25)} = -2.900, p < .008, r_{Y\lambda} = .50$). In both cases, the effect size shows that about 25% of the variance in the posttreatment scores of the EG ($r_{Y\lambda}^2$) was explained by the application of the *EDPL*. In the CG, there was a lack of notable differences, both in primary and in secondary education. With regard to the contrasts of the improvement scores, as a function of the first criterion, significant differences were found for the dimensions D8 and 9 in favor of the EG—in this contrast did not assume equal variances ($t_{(2,58,62)} = -2.458, p < .017, r_{Y\lambda} = .29$). For the second criterion, the contrasts of improvement scores revealed notable differences in favor of the EG, although they did not reach statistical significance either for the global score or for the diverse dimensions. How-

ever, for the third criterion, significant differences were found in favor of the CG for the dimension D2 ($U_{(n=15)} = 3.00, p < .002$). With regard to the diverse subgroups, the contrasts of the diverse dimensions revealed significant differences in the EG for the LD subgroup in the composite dimension (D8 and 9), with a posttreatment gain of 7.02 points ($t_{(2,28)} = -2.486, p < .02, r_{Y\lambda} = .43$). Also for the LD subgroup, significant differences were found in the D1 dimension in the contrasts of the improvement scores, with a mean difference favourable to EG of 8.16 points ($t_{(2,60)} = -2.083, p < .042, r_{Y\lambda} = .26$). For the rest of the subgroups, no mean significant differences were found.

With regard to the results in the Cattell test, for the diverse criteria adopted, the posttreatment improvements of the EG were very significant. These improvements occurred both in primary education ($t_{(2,29)} = -2.245, p < .033, r_{Y\lambda} = .38$) and in secondary education ($t_{(2,26)} = -2.622, p < .014, r_{Y\lambda} = .46$). The positive results were extensive to the analyses carried out in the diverse subgroups, obtaining very significant improvements in the DIS subgroup ($t_{(2,10)} = -3.482, p < .006, r_{Y\lambda} = .74$) and in the SS subgroup ($t_{(2,35)} = 3.085, p < .004, r_{Y\lambda} = .46$). With regard to the contrasts of the improvement scores, significant differences were found in favor of the EG as a function of the second criterion ($t_{(2,37)} = 3.80, p < .001, r_{Y\lambda} = .52$) and also as a function

of the third criterion ($t_{(2,24)} = 3.71, p < .001, r_{Y\lambda} = .59$). Thus, the contrasts of the improvement scores revealed significant differences for the DIS subgroup in favor of the EG. In these contrasts, nonparametric tests were used because of the pretest differences observed ($U_{(n=26)} = 41.00, p < .031$). The results of these analyses for the CG were inconsistent.

Discussion

The results obtained lead to some issues that should be discussed: What were the action mechanisms of the *EDPL* that allowed us to observe positive effects in some students after their application? Why did improvements occur in cognitive performance with a DA device of specific domains? One peculiarity of our study is the fact that the experimental treatment itself is an assessment procedure. The *EDPL* device is proposed essentially to obtain valid information about improvement in the comprehension process of students with special difficulties; however, the process of *mediated interaction* that takes place during its application involves the deployment of a series of *methodological patterns of dynamic assessment (MPDA)* that may have facilitated the emergence of higher-level skills and strategies. In this sense, the evaluators extracted information about these MPDAs that were seen to be efficacious (Navarro, 2008). However, the preferen-

tial analysis of metacognitive skills also led to focusing the mediator activity on processes and strategies of control and self-regulation of learning. This activity may also have facilitated the emergence of this type of skills in the students during the DA process, and this seems to be reflected in the improvement obtained in the dimensions of the criterion-tests more closely related to metacognition. These processes are precisely the ones that seem more susceptible to efficacious educational action to optimize the cognitive functioning of students with learning difficulties or intellectual disability (Carlson & Wiedl, 2000; García-Sánchez, 2002; Guterman, 2002; Navarro & Mora, 2009; Souvignier & Mokhlesgerami, 2006).

In the introduction, it was noted that some investigations in DA had focused their intervention essentially on these skills and the results obtained confirmed their relevance in the improvement of the cognitive functioning of students with difficulties. In this sense, Carlson and Wiedl (2000) found that, when restructuring cognitive tests by the introducing *detailed feedback* and *verbalization* before, during, and after the task—as key factors in the improvement—precisely the students with metacognitive difficulties significantly improved their performance. The greater efficacy of the assessment model proposed by these authors lies in this kind of dynamic assessment of these skills in the students who presented deficits

in these processes. The relevance of verbalization/mediation could be related to a possible restructuring of the processes of thinking, in the sense of optimizing attentional resources, which could positively affect the consideration of the rules to be followed, the strategies to be employed, or the information to be remembered.

Moreover, a second factor in the verbalization process that may have great impact on improving task performance is that it involves the use of an additional coding system; in effect, when verbalization is introduced in the learning process, a double and conjoint coding of information can occur, which would result in improved learning. In our case, this hypothesis could explain the posttreatment improvements in the intelligence test. Thus, the students who obtained posttest gains could have initially used a preferentially perceptive-visual coding system, which was insufficient to perform the task successfully. Only after the application would the two coding systems have merged; on the one hand, the already present perceptive-visual system; and on the other, a symbolic-linguistic coding system, fruit of the use of verbal self-regulation strategies mediated during the implementation of the device.

In this sense, despite the fact that the application of the device acted essentially on reading processes and, therefore, in a mainly verbal sphere, the work developed

—based especially on metacognitive processes and thus more focused on general problem-solving strategies (although applied to a specific domain) than on skills linked to a specific task— could explain the changes in the performance with tests of non-verbal content. The students would have learned strategies to resolve different kinds of problems because the essence of their learning lies in the application of general elements, such as, for example: (a) more reflection prior to acting; (b) maintaining attention for longer periods of time and with more quality; (c) more control and regulation of activity; or (d) reviewing their activity after finishing the task. Thus, the application of these strategies to a test like the *Cattell* test would have contributed significantly to the explanation of the improvements obtained at posttest.

On the other hand, the design used allowed us to control the main sources of error: age and prior performance in intelligence, personal-social adjustment, and comprehension. The levels of equivalence obtained contributed to control over the main effects of regression to the mean, which otherwise would have cast doubt on the improvements observed in the students who started out with lower scores. However, despite the results in cognitive performance, the device did not generate differential effects in comprehension.

Nevertheless, there is a limitation concerning the higher number

of students in the EG who received special support due to the presence of greater reading difficulties. Despite the initial equivalence in reading performance—at least in primary education— this could have been an important obstacle for a valid group comparison of this aspect. The reason for this is the greater possibilities of favorable short- and medium-term evolution in students who present a lower degree of difficulties (which could have occurred in the CG), in comparison to those who displayed more specific difficulties that required special attention (Campione et al., 1987; Navarro & Mora, 2009).

Despite this, with regard to our hypothesis, it is concluded that the *EDPL* produced significant improvements in the students who presented greater difficulties in functions and specific domains and in general intelligence, and it provided information about the mechanisms that may explain the improvements obtained. In this sense, the educational implications derived from these results—especially with regard to learning self-regulation strategies in students with greater difficulties in comprehension and intelligence— are necessarily related to the identification of the action and mediation mechanisms used by the administrators of the device, which are key elements either to obtain valuable information about the intervention or as a guideline and support that optimized the subject's performance

(Grigorenko, 2009). The results are in accordance with those obtained in investigations mentioned in the *introduction* and that precisely underline the relevance of dynamic assessment procedures in students

with greater learning difficulties. Thus, the data obtained will empirically configure the population that could obtain greater benefits from the application of the device in the school setting.

References

- Aguilera, A., & García, I. (2004). Evaluar interviniendo, intervenir evaluando: una propuesta de actuación ante las dificultades del aprendizaje. *Apuntes de Psicología*, 22(3), 309-322.
- Beckmann, N., Beckmann, J. F., & Elliott, J. G. (2009). Self-confidence and performance goal orientation interactively predict performance in a reasoning test with accuracy feedback. *Learning and Individual Differences*, 19(2), 277-282.
- Botella, J., León, O., San Martín, R., & Barriopedro, M. I. (2001). *Análisis de datos en psicología I*. Madrid: Pirámide.
- Campione, J. C., & Brown, A. L. (1987). Linking dynamic testing with school achievement. In C. S. Lidz (Ed.), *Dynamic assessment: An international approach to evaluating learning potential* (pp. 82-115). New York: Guilford Press.
- Campione, J. C., Brown, A. L., & Ferrara, R. A. (1987). Retraso mental e inteligencia. En R. Sternberg (Ed.), *Inteligencia humana, II* (pp. 599-729). Barcelona: Paidós.
- Carlson, J. S., & Wiedl, K. H. (2000). The validity of dynamic assessment. In C. S. Lidz, & J. G. Elliot (Eds.), *Dynamic assessment: Prevailing models and applications* (pp. 681-712). Greenwich, CT: Elsevier-JAI.
- Cattell, R. B., & Cattell, A. K. S. (1974/2001). *Test de factor «G» (Escala 2 y 3)*. Madrid: TEA.
- Cioffi, G., & Carney, J. (1983). Dynamic assessment of reading disabilities. *Reading Teacher*, 36, 764-768.
- Coll, C., Barberá, E., & Onrubia, J. (2000). La atención a la diversidad en las prácticas de evaluación. *Infancia y Aprendizaje*, 90, 111-132.
- Compton, D., Fuchs, D., Fuchs, L., Bouton, B., Gilbert, J., Barquero, L., et al. (2010). Selecting at-risk first-grade readers for early intervention: Eliminating false positives and exploring the promise of a two-stage gated screening process. *Journal of Educational Psychology*, 102(2), 327-340.
- Dörfler, T., Golke, S., & Artelt, C. (2009). Dynamic assessment and its potential for the assessment of reading competence. *Studies in Educational Evaluation*, 35, 77-82.
- Elliott, J. (2003). Dynamic assessment in educational settings: Realising potential. *Educational Review*, 55(1), 15-32.

- Fernández-Ballesteros, R., & Calero, M. D. (2000). The assessment of learning potential: The EPA instrument. In C. S. Lidz, & J. G. Elliott (Eds.), *Dynamic assessment: Prevailing models and applications* (pp. 293-323). Greenwich, CT: Elsevier-JAI.
- Feuerstein, R. (1996). La teoría de la modificabilidad estructural cognitiva. In S. Molina, & M. Fandos (coords.), *Educación Cognitiva I* (pp. 31-75). Zaragoza: Mira Editores.
- García-Sánchez, J. N. (2002). *Aplicaciones de intervención psicopedagógica*. Madrid: Pirámide.
- Gersten, R., & Dimino, J. (2006). New directions in research RTI (response to intervention): Rethinking special education for students with reading difficulties (yet again). *Reading Research Quarterly*, 41(1), 99-108.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 101(3), 371-395.
- Grigorenko, E. (2009). Dynamic assessment and response to intervention. Two sides of one coin. *Journal of Learning Disabilities*, 42(2), 111-132.
- Grigorenko, E., Sternberg, R. J., & Ehrman, M. E. (2000). A theory-based approach to the measurement of foreign language learning ability: The CANAL-FT theory and test. *The Modern Language Journal*, 84, 390-405.
- Guterman, E. (2002). Toward dynamic assessment of reading: Applying metacognitive awareness guidance to reading assessment tasks. *Journal of Research in Reading*, 25(3), 283-298.
- Guthke, J., & Beckmann, J. (2000). The learning test concept and its application in practice. In C. S. Lidz, & J. G. Elliott (Eds.), *Dynamic assessment: Prevailing models and applications* (pp. 17-69). Greenwich, CT: Elsevier-JAI.
- Hacker, D. (1998). Self-regulated comprehension during normal reading. In D. Hacher, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 165-192). Hillsdale, NJ: Erlbaum.
- Haywood, H. C., & Lidz, C. S. (2007). *Dynamic assessment in practice: Clinical and educational applications*. New York: Cambridge University Press.
- Kaniel, S., Tzuriel, D., Feuerstein, R., Ben-Shachar, N., & Eitan, T. (1991). Dynamic assessment, learning, and transfer abilities of Jewish Ethiopian immigrants to Israel. In R. Feuerstein, P. S. Klein, & A. Tannenbaum (Eds.), *Mediated learning experience* (pp. 179-209). London: Freund.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95(2), 163-182.
- Kintsch, W., & Kintsch, E. (2005). Comprehension. In S. G. Paris, & S. A. Stahl (Eds.), *Children's reading comprehension and assessment* (pp. 71-91). Mahwah, NJ: Erlbaum.
- Kozulin, A. (1998). *Psychological tools. A sociocultural approach to education*. Cambridge, MA: Harvard University Press.
- Kozulin, A., & Garb, E. (2001). *Dynamic assessment of EFL text comprehension of at-risk students*. Paper presented at the 9th Conference of the European Association for Research on Learning and Instruction. Fribourg, Switzerland, 2001.
- Lidz, C. S. (2004). Comment on curriculum-based measurement: Describing

- competence, enhancing outcomes, evaluating treatment effects, and identifying treatment nonresponders by Fuchs and Fuchs. *Journal of Cognitive Education and Psychology*, 4(1), 131-133.
- Lidz, C. S., & Elliott, J. G. (Eds.) (2000). *Dynamic assessment: Prevailing models and applications*. Greenwich, CT: Elsevier-JAI.
- Lukas, J. F., & Santiago, K. (2004). *Evaluación educativa*. Madrid: Alianza.
- Moreno, J., & Saldaña, D. (2005). Use of a computer-assisted program to improve metacognition in persons with severe intellectual disabilities. *Research in Developmental Disabilities*, 26, 341-357.
- Navarro, J. J. (2008). *Evaluación dinámica de procesos implicados en la lectura*. Doctoral dissertation. University of Seville (Spain).
- Navarro, J. J., & Mora, J. (2009). Metacogniciones y comprensión de textos. *Investigación en la Escuela*, 69, 99-113.
- Navarro, J. J., & Mora, J. (2011). Analysis of the implementation of a dynamic assessment device of processes involved in reading with learning-disabled children. *Learning and Individual Differences*, 21(2), 168-175, doi: 10.1016/j.lindif.2010.11.008
- Newman, D., Griffin, P., & Cole, M. (1989). *The construction zone*. New York: Cambridge University Press.
- Pardo, A., & San Martín, R. (1998). *Análisis de datos en psicología II*. Madrid: Pirámide.
- Rapp, D., & Broek, P. (2005). Dynamic text comprehension: An integrative view of reading. *Current Directions in Psychological Science*, 14(5), 276-279.
- Resing, W. C. M. (2000). Assessing the learning potential for inductive reasoning (LIR) in young children. In C. S. Lidz, & J. G. Elliott (Eds.), *Dynamic assessment: Prevailing models and applications* (pp. 224-262). Greenwich, CT: Elsevier-JAI.
- Resing, W., & Elliott, J. G. (2011). Dynamic testing with tangible electronics: Measuring children's change in strategy use with a series completion task. *British Journal of Educational Psychology*, 81, doi: 10.1348/2044-8279.002006
- Resing, W., Tunteler, E., de Jong, F., & Bosma, T. (2009). Dynamic testing in indigenous and ethnic minority children. *Learning and Individual Differences*, 19(4), 445-450.
- Ruiz, U., & Camps, A. (2009). Investigar los géneros discursivos en el proceso educativo. *Revista de Psicodidáctica*, 14(2), 211-228.
- Saldaña, D., Mora, J., & Moreno, J. (2007). Evaluación dinámica en retraso mental severo: descripción de un procedimiento ecológico para la valoración de la metacognición. In J. N. García (coord.), *Dificultades del desarrollo* (pp. 1-17). Madrid: Pirámide.
- Sánchez, E. (1999). El lenguaje escrito y sus dificultades: una visión integradora. In A. Marchesi, C. Coll, & J. Palacios (comps.), *Desarrollo psicológico y educación, 3. Trastornos del desarrollo y necesidades educativas especiales* (pp. 127-155). Madrid: Alianza.
- Schlee, J. (1985). Förderdiagnostik - Eine bessere Konzeption? In R. S. Jäger, R. Horn, & K. Ingenkamp (Comps.), *Tests and trends* (pp. 82-208). Weinheim, Germany: Beltz.
- Souvignier, E., & Moxhlesgerami, J. (2006). Using self-regulation as a framework for implementing strategy instruction to foster reading compre-

- hension. *Learning and Instruction*, 16(1), 57-71.
- Sternberg, R. J., & Grigorenko, E. L. (2002). *Dynamic Testing*. Cambridge: Cambridge University Press.
- Swanson, H. L., & Howard, C. B. (2005). Children with reading disabilities: Does dynamic assessment help in the classification? *Learning Disability Quarterly*, 28(1), 17-34.
- Thurman, S. K., & McGrath, M. C. (2008). Environmentally based assessment practices: Viable alternatives to standardized assessment for assessing emergent literacy skills in young children. *Reading & Writing Quarterly*, 24(1), 7-24.
- Vygotsky, L. S. (1934/1995). *Pensamiento y lenguaje*. Barcelona: Paidós.
- Wells, G. (1999). *Dialogic inquiry: Toward a socio-cultural practice and theory of education*. Cambridge, UK: Cambridge University Press.

Juan José Navarro Hidalgo holds a Ph.D. in Psychology, and is a Hired Professor in the Department of Developmental and Educational Psychology of the University of Seville. His research focuses on the dynamic assessment of the processes involved in reading, metacognition, and learning difficulties.

Joaquín Mora Roche holds a Ph.D. in Psychology, and is a Titled Professor in the Department of Developmental and Educational Psychology of the University of Seville. He directs the research group *Cognitive enrichment and special educational needs*. His research revolves about disability, metacognition, and language. He is the author of the cognitive enrichment program *Comprehending and Transforming*.

Received date: 11-11-11

Review date: 23-3-11

Accepted date: 22-6-11

ANNEX I

**Extract of the session
corresponding to the application
group (B), Public Primary and
Secondary Center Antonio
Machado (teacher María José
Alés Hervás)**

Transcription of the mediation carried out with C., a male student from 5th grade who presents special difficulties in the global comprehension of texts. The mediation is based on activity 22: «*Use of the structure of a text*», which belongs to the block of textual integration processes. This activity basically consisted of the student reading from a set of texts, and then relating their content. Two of the texts

were narrative in structure, one of them respected the logical structure and the other one was disordered in this sense. The other two texts were expositive and, as with the narratives, one of them was logically ordered and the other was not. They were simple texts with familiar contents; the difficulty of the task should not lie in the complexity of the text contents. We wished to assess whether the students identified the textual structure and whether they used structural elements to access the global comprehension of the texts. We also wished to assess the structuring of the text summary carried out by the student, that is, whether he/she wrote a text with defined meaning and structure.

Mediation performed	Analysis of the methodological patterns of dynamic assessment used in the mediation
<p><i>The whales (ordered text)</i></p> <p>Starts to read and goes back because he says he could not understand.</p> <p><i>C.: I had to start three times and it was his fault! (His classmate, who does not focus on the activity). He activates his prior knowledge when he finds a word he had seen in another text.</i></p> <p><i>C.: Filter! Like the layer that surrounds the earth and filters the sun.</i></p> <p>He asks the words he does not understand or know. He controls the activity, he knows that he must understand and he introduces latencies to make sure. He rereads the text, without needing external mediation. He is supervising the activity. He explains the content of the text coherently, with the proper sequence. There is no external verbal mediation by him.</p>	<p>The Assessor takes notes about the implementation of self-regulation processes (supervision strategies), processes of personal-social adjustment, and processes of integration of text-prior knowledge. The student relates one of the words of the new text to the content of the previous text; however, in the transcription, there is no recording by the Assessor of a positive reinforcement response or of confirmation of the response, or of the subject's attitude. She makes notes about the implementation of processes of self-regulation (supervision strategies). She also makes notes about the quantity and quality of help needed to accomplish the activity. *The Assessor asks questions to facilitate the subject's self-assessment with regard to whether he liked the reading, and the ease or difficulty of his comprehension. *She</p>

Mediation performed	Analysis of the methodological patterns of dynamic assessment used in the mediation
<p>Assessor: <i>Did you like reading?</i> C.: <i>Yes</i> Assessor: <i>Was it easy or difficult?</i> C.: <i>Easy or regular.</i> Assessor: <i>But did this text say the things correctly?</i> C.: <i>Yes, because in the other text, the one of Rabin, the things were wrong. (We look for the disordered text to remember it).</i> Assessor: <i>And when it was disordered, what happened, was it easier to understand?</i> C.: <i>No, it was more difficult because you get mixed up.</i></p> <p><i>Breathing (disordered text)</i></p> <p>C.: <i>Why is the title down here?</i> Assessor: <i>Where should it be?</i> C.: <i>Up.</i> He reads very fast. C.: <i>Ready! But I'm going to read it again.</i> While reading, he comments on the content of the text. Assessor: <i>Did you understand all the words of the text?</i> He asks several words he didn't know, but which he had not asked spontaneously. When he finishes reading: C.: <i>This part can't be the first part because it says: «secondly...» Then there has to be another one before.</i> Assessor: <i>Good, and what does the text tell you?</i> C.: <i>I don't know, it's very difficult.</i> Assessor: <i>What does it talk about? Whales?</i> C.: <i>No, about lungs.</i> Assessor: <i>What else?</i> C.: <i>It's very difficult because it talks about lungs, oxygen...</i> We reread it, paragraph by paragraph, trying to understand the global meaning of the text. It takes a lot of effort but at last, he gets a slight idea of the text. Incoherencies in the structure of the text were detected.</p>	<p>relates the text he is working on to the previous one in order to provoke the subject's assessment about the structural characteristics of both texts. *She facilitates and accompanies the student's activity, searching for the disordered text so as to facilitate comparison. *The Assessor establishes gradual support by means of questions that guide the subject toward a conclusion.</p> <p>The student underlines an incoherence in the text about its structure, and the Assessor attempts to *deepen and question the subject's comments with a new question. She makes notes about the subject's behavior regarding the implementation of processes of self-regulation (supervision). *The Assessor asks a question that attempts to provoke a process of reflection and self-assessment in the student with regard to his knowledge of the vocabulary of the text. *She asks the subject about the contents of the text, *establishing gradual support by means of new questions when she sees the subject's difficulty to tell her what the text is about. *She sequences the content of the activity to facilitate its accomplishment. *The Assessor reinforces and confirms the subject's response.</p>

NOTES USED IN THE TRANSCRIPTION:

The notes in boldface reflect the assessor's reflections or comments about the process of mediation or about the student's response.

In the column corresponding to the analysis carried out, asterisks indicate references to methodological patterns of dynamic assessment (MPDA).

Evaluación dinámica de las dificultades de aprendizaje de la lectura

Dynamic Assessment of Reading Difficulties

Juan-José Navarro y Joaquín Mora

Universidad de Sevilla¹

Resumen

En este artículo se recogen los cambios producidos en un grupo de estudiantes con dificultades de comprensión lectora (DL) como respuesta a la aplicación experimental de un dispositivo de evaluación dinámica de procesos implicados en la lectura (EDPL). La muestra está integrada por 133 estudiantes de 9 a 16 años con DL (60 en el grupo experimental y 73 en el grupo control). Se utiliza un diseño pretest/postest para la evaluación de los resultados. Los análisis efectuados permiten obtener información diferenciada sobre el impacto de EDPL en función del grado de dificultad previamente establecido, configurando empíricamente la población que obtendría mayores beneficios de su aplicación en el contexto escolar. Los resultados muestran mejoras significativas en comprensión, ajuste personal-social e inteligencia en aquellos estudiantes que presentan mayores dificultades, tanto en dominios específicos como en inteligencia. Asimismo, los análisis otorgan información acerca de los mecanismos que podrían explicar las mejoras.

Palabras clave: Evaluación dinámica, comprensión lectora, dificultades de aprendizaje de la lectura.

Abstract

This article describe the changes produced in a group of students with reading difficulties (RD) in response to the experimental application of a dynamic assessment device of processes involved in reading (EDPL). The sample consists of 133 students of 9-16 years with RD (60 in the experimental group and 73 in the control group). The design is a pretest/posttest for impact evaluation. Analyses allow differentiated information on the impact of EDPL device depending on the degree of difficulty previously established; empirically setting the population would get more benefits from its application in the school context. The results show significant improvements in reading comprehension, personal-social adjustment and intelligence in those students who have greater difficulties both in specific domains like intelligence. Likewise, the analysis gives information about the mechanisms that might explain the improvements.

Keywords: Dynamic assessment, reading comprehension, reading difficulties.

Correspondencia: Juan José Navarro Hidalgo, Departamento de Psicología Evolutiva y de la Educación, Universidad de Sevilla, c/ Pirotecnia, s/n., 41013, Sevilla. E-mail: jnnavarro@us.es.

¹ Deseamos agradecer la colaboración inestimable del profesorado perteneciente a los centros educativos participantes, así como la del CEP de Alcalá de Guadaíra (Sevilla). Asimismo, agradecemos a la Dirección General de Innovación Educativa y Formación del Profesorado de la Junta de Andalucía por el apoyo prestado a través de la concesión de una licencia de estudios al primer autor.

Introducción

Los modelos dinámicos de evaluación se apoyan en la interacción que mantienen los sujetos evaluados con otros más expertos, que prestan su ayuda estableciendo un proceso de mediación entre competencias actuales y exigencias de la tarea. De este modo, perfilan un espacio de construcción del conocimiento especialmente sensible a la acción pedagógica, facilitando de este modo la identificación del tipo y grado de ayuda educativa necesaria, y estableciendo una medida dinámica de los procesos evaluados en contraposición a medidas de corte más estático, que informarían esencialmente sobre habilidades ya consolidadas (Elliott, 2003; Grigorenko, 2009; Newman, Griffin, y Cole, 1989; Sternberg y Grigorenko, 2002; Vygotsky, 1934/1995). La relevancia de la aplicación de estos modelos de evaluación parece ser mayor ante estudiantes que presentan dificultades de aprendizaje o discapacidad cognitiva (Campione y Brown, 1987; Carlson y Wiedl, 2000; Guthke y Beckmann, 2000; Saldaña, Mora, y Moreno, 2007; Swanson y Howard, 2005). Por otra parte, la consideración del contexto social caracterizado por la interacción y el apoyo en el que se desenvuelve habitualmente el proceso educativo, puede contribuir significativamente al incremento de la validez ecológica de estos procedimientos, facilitando en este sentido su aplicación y generalización en el contexto escolar

(Schlee, 1985; Thurman y McGrath, 2008; Wells, 1999).

Una de las razones fundamentales que se esgrimen a favor de la evaluación dinámica (ED) es su capacidad para identificar diferencias interindividuales minimizando la influencia de la historia previa de los estudiantes (Kozulin, 2000; Resing, Tunteler, de Jong, y Bosma, 2009). En efecto, al tratarse de actividades que tratan de valorar el proceso y la capacidad de aprendizaje de determinados conceptos o habilidades, estableciendo un punto de partida común con respecto a los contenidos a trabajar, y centrando su atención especialmente en habilidades y competencias *en desarrollo*, estos procedimientos estarían indudablemente mejor situados que las pruebas convencionales (Dörfler, Golke, y Artelt, 2009; Lidz y Elliot, 2000; Sternberg y Grigorenko, 2002). Algunos estudios han puesto de manifiesto que niños con muy diferentes historias de escolarización, podían igualar su rendimiento en pruebas que evaluaban su potencial de aprendizaje (Feuerstein, 1996; Kaniel, Tzuriel, Feuerstein, Ben-Shachar, y Eitan, 1991). También hay evidencias de que estudiantes con especiales dificultades mejoraban significativamente su rendimiento en situaciones de aprendizaje asistido o en colaboración, así como tras la aplicación de procedimientos dinámicos de evaluación, mostrando estos trabajos en general efectos positivos en los grupos de tratamiento (Campione, Brown, y

Ferrara, 1987; Fernández-Ballesteros y Calero, 2000; Guthke y Beckmann, 2000; Resing, 2000; Resing et al., 2009; Resing y Elliott, 2011; Saldaña, Mora, y Moreno, 2007). En este sentido, se ha informado acerca de mejoras significativas en dominios específicos, como la memoria de trabajo, la aritmética o la lectura (Swanson y Howard, 2005), así como en habilidades generales medidas con pruebas de CI tras aplicaciones centradas esencialmente en procesos metacognitivos, en los que se manipulaban las condiciones de aplicación de las tareas con relación a la cantidad o el tipo de verbalización y retroalimentación que recibía el sujeto durante la mediación (Beckmann, Beckmann, y Elliott, 2009; Carlson y Wiedl, 2000).

Por otra parte, la ED permite acceder a la valoración integrada de distintos procesos implicados en una actividad (Rapp y Broek, 2005). En este sentido, la información sobre los mecanismos a través de los cuales un alumno construye y optimiza su aprendizaje puede ser significativa para orientar la acción educativa (Ruiz y Camps, 2009), permitiendo de esta forma establecer conexiones reales entre los procesos de evaluación e intervención, aspecto considerado clave para mejorar los procesos educativos (Aguilera y García, 2004; Coll, Barberá, y Onrubia, 2000; Gersten y Dimino, 2006; Grigorenko, 2009). En sus planteamientos sobre evaluación, Vygotsky hizo una distinción particular entre la ED de funciones cognitivas generales

(vertiente ésta más desarrollada) y la ED de dominios específicos de aprendizaje escolar, como la lectura o la aritmética (Kozulin y Garb, 2001). La naturaleza de las tareas propuestas desde cada uno de estos enfoques tiene, a nuestro entender, una relevancia fundamental en educación. En este sentido, por una parte estarían las pruebas dinámicas que incluyen *tareas nuevas para el alumno*; estas tareas se caracterizan por su descontextualización y su semejanza con los test convencionales, y pretenden evaluar la capacidad de aprendizaje (Kozulin, 2000; Sternberg y Grigorenko, 2002). Entre sus ventajas se encontrarían principalmente: (a) evaluar el potencial de aprendizaje en tareas «libres de cultura» y (b) una supuesta igualdad de condiciones al enfrentarse a las tareas. Entre los inconvenientes podríamos destacar: (a) evaluar el potencial de aprendizaje «en abstracto», sin correspondencia con los aprendizajes significativos para la escuela, (b) la descontextualización de la actividad de aprendizaje, (c) la escasez/ausencia de validez ecológica y (d) un supuesto menor poder predictivo del rendimiento como consecuencia de los elementos citados. Por otra parte, estarían aquellas pruebas que incluyen *tareas con contenidos relacionados con los aprendizajes escolares* (Cioffi y Carney, 1983; Dörfler et al., 2009; Grigorenko, Sternberg, y Ehrman, 2000; Haywood y Lidz, 2007; Sternberg y Grigorenko, 2002). Entre sus ventajas contaríamos: (a) la contex-

tualización de los instrumentos de evaluación, (b) una mayor validez ecológica de los resultados, y (c) un supuesto mayor poder de predicción del rendimiento escolar. Entre sus inconvenientes: (a) que al ser actividades cuyo contenido está saturado de información que puede formar ya parte del bagaje de conocimientos del sujeto, sería más difícil establecer la capacidad de aprendizaje como tal, y, como consecuencia de lo anterior, (b) la necesidad de establecer exhaustivamente el nivel inicial de desarrollo para poder determinar las mejoras, así como la necesidad, en definitiva, de construir dispositivos de evaluación específicos, sensibles a los procesos implicados y eficaces en el contexto escolar (Lidz, 2004; Thurman y McGrath, 2008).

Objetivos e hipótesis de investigación

El presente trabajo pretende analizar los efectos producidos en un grupo de estudiantes con dificultades de comprensión lectora (DL) a raíz de la aplicación experimental de un dispositivo de evaluación dinámica de procesos implicados en la lectura (EDPL). Nuestra hipótesis contemplaba posibles mejoras en comprensión, así como en los niveles de ajuste personal-social y rendimiento cognitivo. Esta hipótesis se fundamentaba en los procesos de *interacción mediada* que tienen lugar durante la aplicación, así como en el contenido de carácter meta-

cognitivo que está en la base de la mayoría de las pautas de mediación que ofrecen los aplicadores durante las sesiones.

Método

Participantes

En la investigación participaron 13 centros educativos (7 de primaria y 6 de secundaria obligatoria —ESO—) localizados en Sevilla y Cádiz (Andalucía-España). La muestra estuvo integrada por 133 estudiantes con DL, de los cuales 60 (23 chicas y 37 chicos) formaron parte del grupo experimental (GE) y 73 (29 chicas y 44 chicos) del grupo control (GC). El rango de edad se situaba en 9-16 años. La asignación de los estudiantes a estos grupos se realizó en función de las posibilidades de participación del profesorado colaborador en el Grupo de Trabajo (GT) que implementó el dispositivo *EDPL*. Doce profesores/orientadores (aplicadores) integraron este GT y cinco integraron el GC. El menor número de integrantes de este grupo se explica debido a que algunos orientadores participantes del GC atendían a grupos amplios de estudiantes y, así mismo, algunos orientadores del GE desarrollaban su trabajo en centros que contaron con estudiantes que finalmente conformaron el GC. En ambos casos, se trataba de profesionales implicados en proyectos de innovación, con formación

y experiencia en educación especial, psicología o pedagogía. Por su parte, 48 profesores-tutores colaboraron en la cumplimentación de escalas de valoración y 6 profesores ajenos a los centros participaron en la aplicación de las pruebas-criterio. La selección inicial de la muestra se realizó con base en la información proporcionada mediante una plantilla por el profesorado especialista y orientadores acerca de la presencia significativa de DL. Se contempla-

ron asimismo diversos subgrupos en función de la información otorgada por los equipos de orientación educativa y por los centros, y que estaban representados en cada una de las condiciones experimentales. Estos subgrupos no eran excluyentes y su consideración en los análisis posteriores debía contribuir a delimitar mejor el ámbito de eficacia del dispositivo *EDPL*. El lector puede encontrar más detalles de la muestra en la tabla 1.

Tabla 1

Composición y Distribución de la Muestra

Distribución por etapa y grupos de participación				Profesorado participante		Subgrupos contemplados para el análisis	GE	GC
Etapa	Estudiantes	GE	GC	GE (integrantes del GT)	12	Dificultades de aprendizaje	31	47
Primaria E.S.O.	68	32	36	GC	5	Discapacidad cognitiva	13	16
	65	28	37	Validación expertos	4	Deprivación socio-cultural	16	10
				Aplicación pruebas-criterio	6	<i>Recibe apoyo específico</i>	39	26
				Profesorado colaborador	48			
Totales	133	60	73					

Nota: El subgrupo *Recibe apoyo específico* estaba compuesto por estudiantes con DL que recibían atención educativa especializada, estando integrado por sujetos de los otros subgrupos.

Instrumentos**El dispositivo *EDPL***

EDPL (Navarro, 2008) es un conjunto de estrategias y actividades de ED a raíz de las cuales se pretende obtener información acerca del proceso de resolución de las tareas propuestas, así como de los mecanismos que contribuyen a la

mejora durante la actividad. La evaluación de la lectura propuesta, en consonancia con el marco teórico de referencia adoptado, caracterizado básicamente por considerar la lectura como un proceso interactivo (Graesser, Singer, y Trabasso, 1994; Rapp y Broek, 2005), sistémico y dinámico que se construye socialmente en contextos de interacción (Wells, 1999), está basada en

un proceso de *interacción mediada* entre evaluador y estudiante con dificultades, proporcionando una conexión entre evaluación e intervención (Aguilera y García, 2004; Grigorenko, 2009). En este sentido, los evaluadores asumen una posición activa en las situaciones de evaluación, modifican sus preguntas e incluso los propios instrumentos en función de la respuesta del sujeto; ofrecen herramientas o pautas

de control y regulación de la actividad para observar y valorar el grado de incorporación de éstas en la resolución de la tarea; intervienen en el proceso mediando su realización, con el objetivo de evaluar el proceso mismo de ejecución y el grado y tipo de apoyo que necesita el sujeto (Anexo I).

El dispositivo *EDPL* se compone de 32 actividades de evaluación (figura 1) encuadradas en una

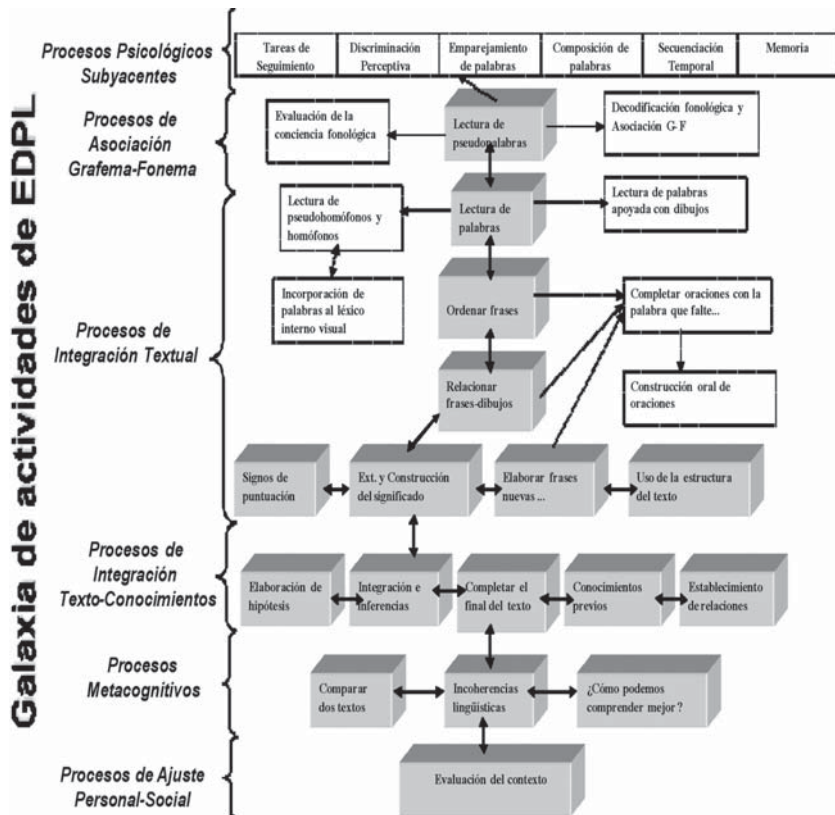


Figura 1. Actividades de evaluación de los distintos procesos de lectura contemplados en EDPL (tomado de Navarro y Mora, 2011).

serie de procesos que han sido contemplados en función de la investigación precedente (Compton et al., 2010; Graesser, Singer, y Trabasso, 1994; Hacker, 1998; Kintsch, 1988; Kintsch & Kintsch, 2005; Sánchez, 1999). Por su parte, cada actividad recoge: (a) el proceso que se pretende evaluar; (b) descripción de la actividad a realizar; (c) metodología propuesta para su aplicación; (d) patrones de mediación y evaluación de procesos metacognitivos, es decir, las herramientas y pautas de control y regulación de la actividad a las que hacíamos referencia anteriormente; y (e) criterios o indicadores de evaluación.

Para facilitar la recogida y valoración de la información, los aplica-

dores disponían de *Hojas de registro y evaluación* de cada actividad. En estas hojas debían plasmar el proceso de mediación y evaluación efectuado. Esta evaluación implicaba analizar el proceso de aprendizaje y resolución de las tareas, lo que conllevaba dificultades relativas a la «invisibilidad» de los procesos cognitivos y emocionales puestos en juego, debiendo necesariamente inferir el evaluador su puesta en práctica. En este sentido, se elaboró un listado de indicadores (tabla 2) que facilitarían la tarea de evaluar con mayor precisión la presencia y calidad de los distintos procesos contemplados. Estos indicadores fueron entregados a los aplicadores formando parte de cada

Tabla 2

Indicadores del Bloque de Procesos de Integración Texto-Conocimientos Previos

Integra nuevas ideas con conocimientos y experiencias anteriores
Utiliza conocimientos y experiencias anteriores para construir significados
Es sensible al estilo, los presupuestos, la perspectiva y las pretensiones del autor/a y hasta puede llegar a cuestionar estos aspectos
Expresa opiniones, juicios o intuiciones acerca de los contenidos del texto
Establece relaciones entre diversos textos
Formula y responde preguntas acerca del texto
Formula hipótesis. Anticipa resultados
Contrasta y verifica las hipótesis
Realiza inferencias relativas al texto y que lo sobrepasan
Revisa las ideas del texto desde lo que ya sabe

una de las actividades y debían ser valorados en las Hojas de registro y evaluación. Se desarrolló un sistema de análisis que cuantificara las valoraciones efectuadas (Lukas y Santiago, 2004). Esto nos permitiría operativizar el proceso de aplicación y disponer de *puntuaciones dinámicas* para cada actividad y también globales. El sistema adoptado (Moreno y Saldaña, 2005), tenía como base una escala con carácter gradual compuesta por cuatro apreciaciones: (1) No aparece la conducta. El indicador no se ha puesto en práctica; (2) Indicios, conducta rudimentaria; (3) Conducta claramente apreciable (aunque no sea de

gran calidad); y, (4) Conducta intensa o de gran calidad.

Batería de criterio

- *Test de factor «G» de Cattell, Escala 2-Forma A* (Cattell y Cattell, 1974/2001).
- *Escala de Ajuste Personal-Social con relación a la Lectura, APSL*. Esta escala, de elaboración propia (Navarro, 2008), otorga información acerca de 9 dimensiones (tabla 3) y se compone de ochenta ítems presentados al alumnado en una estructura de frases cortas ante las que éste debe mostrar individualmente su

Tabla 3

Dimensiones de la Escala APSL

-
- D1 Conocimiento de las propias limitaciones cognitivas actuales, potencialidades y posibilidades de aprendizaje, dificultades y vías de solución.
 - D2 Conocimiento de las propias expectativas, autoconcepto académico y estilo atribucional.
 - D3 Conocimiento de la propia actitud ante la lectura: emociones, sentimiento de rechazo vs satisfacción y gusto por la lectura, tolerancia a la frustración, persistencia ante la dificultad, autonomía vs dependencia.
 - D4 Conocimiento de los propios intereses y motivaciones.
 - D5 Percepción del apoyo que ofrece el contexto escolar y familiar, y de las condiciones en las que se desenvuelve.
 - D6 Percepción de la relevancia del tema: la lectura, para qué sirve, cuál es su importancia.
 - D7 Conocimiento de las propias estrategias de relación con los demás en tareas de lectura.
 - D8 Conocimiento de las propias estrategias de control y regulación consciente del proceso de lectura (planificación, supervisión, autoevaluación).
 - D9 Reconocimiento de la capacidad de generalización de estos conocimientos o estrategias.
-

acuerdo o desacuerdo. Se obtiene una puntuación global y puntuaciones relativas a las distintas dimensiones.

- *Comprensión de textos*: Se utilizaron dos pruebas de evaluación de la comprensión de textos: *ECO1*, para 2.º y 3.º ciclo de primaria, y *ECOS*, de aplicación en E.S.O (tabla 4). Fueron elaboradas en función de los objetivos del estudio (Navarro, 2008) y pretenden dar respuesta a la valoración del conocimiento y uso de estrategias lectoras, así como de la comprensión global. *ECO1* y *ECOS* se componen de 10 textos que presentan diversas

estructuras y pertenecen a diversos ámbitos relacionados con las áreas curriculares de primaria y secundaria respectivamente, seguidos de una serie de preguntas (43 para *ECO1* y 56 para *ECOS*) con opción múltiple de respuesta. La aplicación de estas pruebas otorga una puntuación global en comprensión, al tiempo que puntuaciones relativas a las distintas dimensiones contempladas.

Debemos destacar que los análisis sobre la validez y fiabilidad de las pruebas de comprensión y ajuste personal-social, así como la validación de expertos realizada previa-

Tabla 4

Dimensiones en las Pruebas de Comprensión de Textos ECO1 y ECOS

Concepción sobre la lectura: Para qué sirve leer, en qué consiste y que implica la lectura.

Conocimiento de estrategias que pueden aplicarse: estrategias de autorregulación (planificación, Supervisión, autoevaluación), transferencia y generalización, integración textual e integración texto-conocimientos previos.

Conocimientos previos (sólo en ECOS): Conocimientos sobre la temática del texto, situaciones planteadas y esquemas de acción descritos.

Integración textual: acceso al significado de palabras, integración lineal de las ideas, integración de todas las ideas en un esquema, construcción de proposiciones globales, distinción de la información relevante, jerarquización y secuenciación de la información, acceso a la información contenida en cuadros, tablas y esquemas, identificación y uso de la estructura textual.

Integración del texto con los conocimientos y experiencias previas: establecimiento de relaciones, formulación de hipótesis, realización de inferencias, intención del autor, revisión crítica del texto desde lo ya conocido.

Supervisión del proceso de comprensión: detección de incoherencias y superación de las dificultades mediante la aplicación de estrategias de procesamiento léxico, sintáctico y semántico.

mente a su aplicación, arrojaron resultados positivos. Se realizó una aplicación de estas pruebas a grupos-clase completos pertenecientes a los centros colaboradores, obteniendo entre otros los siguientes datos: prueba *ECOI* ($n = 179$): coeficiente de fiabilidad «alfa» = .81; prueba «Kolmogorov-Smirnov» de normalidad = .088 ($p < .002$); índice medio de homogeneidad (IMH) = .32; índice de dificultad medio (IDM) = .65; prueba *ECOS* ($n = 43$): «alfa» = .89; prueba «Shapiro-Wilk» de normalidad ($n < 50$) = .94 ($p < .027$); IMH = .37; IDM = .51; escala *APSL* ($n = 98$): «alfa» = .88; «Kolmogorov-Smirnov» = .077 ($p < .179$);

IMH = .32; IDM = .31. Es necesario señalar la ausencia de normalidad en los datos de *APSL*, lo que implicó el uso de pruebas no paramétricas en algunos análisis efectuados con muestras reducidas.

Diseño y procedimiento

Para la evaluación de impacto se utilizó un diseño *pretest-postest con grupo control*. En la tabla 5 hemos sintetizado el procedimiento seguido.

La aplicación de *EDPL* fue realizada en pequeño grupo. Los 12 aplicadores recibieron formación relativa a las bases teórico/metodológicas de la propuesta. Estas sesio-

Tabla 5

Procedimiento Seguido en la Investigación

-
1. Aplicación piloto de instrumentos diseñados y validación de expertos
 2. Contactamos con colaboradores en la fase experimental y con las familias
 3. Programación y comienzo de las sesiones del Grupo de Trabajo (GT). Formación de aplicadores
 4. Formación de los colaboradores que efectuarían la pasación de pruebas-criterio
 5. Pasación de pruebas-criterio a GE y GC
 6. Aplicación del dispositivo en el GE por parte de los profesores/aplicadores / El GC recibe *otros tratamientos*
 7. Seguimiento y evaluación del proceso de aplicación de *EDPL* por parte del GT
 8. Aplicación postest de las pruebas-criterio
 9. Sesiones de evaluación en el GT del dispositivo *EDPL* y de su aplicación
 10. Recogida final de datos. Determinación de resultados en función de hipótesis
-

Tabla 6

Análisis de Equivalencia Inicial Global y por Subgrupos entre GC y GE para las Variables Inteligencia, Ajuste y Comprensión

Medidas pretest	Global		Subgrupos			
			DA	DIS	DEP	AE
Inteligencia	M-GC (DT)	80.91 (18.05)	87.71 (16.36)	68.25 (14.30)	88.80 (16.83)	70.70 (14.61)
	M-GE (DT)	82.57 (14.00)	83.06 (15.03)	76.31 (10.76)	86.70 (13.20)	79.82 (13.37)
	Levene	3.78	.45	.53	.59	0.18
	Sig.	.054	.50	.47	.45	0.67
	<i>t</i>	(2, 103) = .53	(2, 53) = -1.09	(2, 27) = 1.68	(2, 19) = -2.29	(2, 57) = 2.40
	Sig.	.60	.28	.11	.77	.02*
Ajuste personal-social	M-GC (DT)	95.06 (15.90)	95.50 (15.29)	97.38 (11.47)	89.28 (24.02)	92.65 (13.36)
	M-GE (DT)	97.98 (13.91)	95.78 (14.32)	97.76 (12.40)	102.13 (14.31)	94.77 (13.18)
	Levene	.16	.05	.24	3.44	.30
	Sig.	.69	.82	.63	.08	.59
	<i>t</i>	(2, 115) = .98	(2, 63) = .08	(2, 27) = .09	(2, 21) = 1.62	(2, 62) = .63
	Sig.	.33	.94	.93	.12	.53
Comprensión (primaria)	M-GC (DT)	86.16 (8.91)	88.75 (8.55)	80.75 (9.04)	82.66 (8.22)	80.47 (7.72)
	M-GE (DT)	84.18 (8.75)	86.92 (8.57)	84.35 (6.78)	81.95 (9.64)	84.79 (9.53)
	Levene	.25	.31	.15	.26	.13
	Sig.	.62	.58	.71	.61	.72
	<i>t</i>	(2, 66) = -.93	(2, 31) = -.58	(2, 9) = .76	(2, 22) = -.19	(2, 26) = 1.14
	Sig.	.36	.36	.36	.36	.36
Comprensión (secundaria)	M-GC (DT)	91.17 (8.02)	92.42 (8.20)	88.56 (7.28)	82.41 (4.09)	88.20 (6.75)
	M-GE (DT)	88.25 (10.24)	90.73 (9.68)	79.52 (6.34)	83.14 (12.30)	83.33 (5.67)
	Levene	1.67	.95	1.07	1.59	1.52
	Sig.	.20	.34	.32	.28	.23
	<i>t</i>	(2, 63) = -1.29	(2, 43) = -.64	(2, 16) = -2.59	(2, 4) = .08	(2, 35) = -2.38
	Sig.	.20	.53	.02*	.94	.02*
Presencia de subgrupos en GE y GC	GC		47	16	10	26
	GE		31	13	16	39
	χ^2		3.28	.31	1.39	2.60
	Sig.		.070	.58	.24	.11

Nota: En los contrastes se han asumido varianzas iguales; χ^2 = Chi cuadrado; M-GC/GE (DT) = Media del grupo control/experimental y desviación típica; DA = Dificultades de aprendizaje (excluyendo deprivación y discapacidad); DIS = Discapacidad cognitiva; DEP = Deprivación sociocultural (excluyendo discapacidad); AE = Recibe apoyo específico; * $p < .05$.

nes de formación tuvieron lugar en el marco de un GT coordinado por el investigador principal y a través del cual se realizó el seguimiento y la evaluación de todo el proceso. El contenido específico de estas sesiones incluía la presentación, discusión y análisis conjunto de las actividades propuestas, la ejemplificación de la aplicación de los patrones de mediación recogidos en el dispositivo, el análisis de vídeos y el seguimiento de las sesiones iniciales. El tiempo de aplicación fue de 16 semanas; cada grupo de aplicación (11 en total) invirtió un promedio de 2 sesiones semanales de 45 minutos. En función de los objetivos propuestos, los aplicadores debían realizar, en una secuencia ya pre-determinada, las actividades que se encuentran en los rectángulos tridimensionales de la Figura 1. Por su parte, dado que el GC estaba integrado por estudiantes con DL que recibían apoyo escolar, se trataba en realidad de un grupo de *otros tratamientos*, cuyo efecto, en cualquier caso, quedaría aleatorizado en función de la diversidad de actuaciones, centros y docentes participantes.

En la tabla 6 recogemos los valores de equivalencia inicial entre GE y GC para diversas variables, incluida la presencia en ambos de sujetos pertenecientes a los diversos subgrupos. En los análisis no hallamos generalmente diferencias significativas inter-grupos. Asimismo, los valores de la prueba «Levene» de homogeneidad de varianzas mostraron esta equivalencia de partida.

En los contrastes relativos a la inteligencia, el ajuste y la comprensión fue aplicada la prueba *t-Student* para grupos independientes. Con relación a la edad, los valores promedio del GE (12.40, DT = 2.44) y del GC (12.05, DT = 2.34), así como la prueba «Levene» y la *t-Student*, mostraron equivalencia ($F = .59$, $p < .44$; $t_{(2, 131)} = .83$, $p < .41$). Por su parte, el sexo también mostró equivalencia inter-grupos (χ^2 GE/GC niñas = .69; $p < .41$; χ^2 GE/GC niños = .61; $p < .44$).

Análisis de datos y tratamiento estadístico

Además de la designación previa del profesorado como criterio para establecer las DL, en nuestros análisis fueron consideradas las puntuaciones iniciales en las pruebas de comprensión, ajuste y CI, estableciéndose tres criterios de análisis en función de la valoración de las dificultades: (a) Valoración inicial del profesorado como alumno con DL; (b) Valoración del profesorado + puntuación inicial en Comprensión [o Ajuste] ($\leq 84/100$); y (c) Valoración del profesorado + puntuación inicial en Comprensión [o Ajuste] ($\leq 84/100$) + puntuación inicial en CI ($\leq 84/100$). En este sentido, un alumno podía cumplir un solo criterio, dos, o los tres. De este modo, ha sido posible analizar los resultados desde una óptica complementaria en función del criterio de análisis adoptado, permitiendo asi-

mismo obtener información diferenciada sobre el impacto de *EDPL* en función del grado de dificultad del estudiante. Por otra parte, todas las puntuaciones manejadas en la investigación fueron transformadas a través de la aplicación de una *escala* derivada con media 100 y DT 16 (Botella, León, San Martín, y Barriopedro, 2001). Para los contrastes con la escala *APSL* en subgrupos pequeños utilizamos las pruebas no paramétricas de «Wilcoxon» para muestras relacionadas y «U de Mann-Whitney» para muestras independientes. Asimismo, procedimos a calcular los valores de tamaño del efecto para cada uno de los contrastes efectuados mediante el coeficiente de correlación *biserial-puntual* $r_{Y\lambda}$ de Pearson (Pardo y San Martín, 1998). Estos valores permiten obtener información no sólo sobre la existencia o no de un efecto significativo del tratamiento aplicado, sino sobre el impacto que éste ha tenido sobre las variables estudiadas. En este sentido, la estandarización propuesta asume que aquellos valores de $r_{Y\lambda}$ que se sitúen por encima de .37 permitirían hablar de un tamaño grande del efecto y, por tanto, de un impacto significativo del tratamiento aplicado.

Resultados

Los análisis han sido efectuados, para cada una de las pruebas, en función de los tres criterios de análisis establecidos (tabla 7), y teniendo

en cuenta asimismo los distintos subgrupos contemplados (tabla 8). En la presentación de los resultados iremos haciendo referencia a los efectos más relevantes observados en cada prueba-criterio, incluyendo algunos datos provenientes de los contrastes efectuados para las distintas dimensiones que integran.

En primer lugar debemos admitir que el dispositivo no tuvo efectos diferenciales significativos sobre el rendimiento en comprensión en función del criterio de designación previa del profesorado, es decir, cuando fueron contemplados todos los sujetos. Aunque la inspección de los resultados muestra una evolución positiva en las ganancias del GE en *ECO1* a medida que el criterio de análisis implicaba mayores dificultades, los efectos se ven ciertamente ensombrecidos a tenor de los resultados del GC. A pesar de la equivalencia inicial inter-grupos, la mayor presencia en el GE de alumnos que recibían apoyo específico (AE), podría haber tenido una influencia significativa en la obtención de bajas puntuaciones postest en este grupo. En la *discusión* volveremos sobre este asunto. Por otra parte, hemos contrastado, mediante la prueba *t* para muestras independientes, las puntuaciones de mejora (PM) para *ECO1* entre GE y GC. Las diferencias promedio son favorables al GC, llegando a ser significativas para la dimensión *Integración textual* en función del primer criterio de análisis ($t_{(2,64)} = 2.190, p < .032$) y del segundo ($t_{(2,22)} = 2.147, p < .043$).

Tabla 7

Resultados en GC y GE para la Puntuación de Mejora (Postest-Pretest) en las Distintas Pruebas, en Función de los Tres Criterios de Análisis Establecidos

Pruebas		(a) Designación previa del profesorado			(b) Criterio del profesorado + puntuación inicial en comprensión/ajuste (≤ 84)			(c) Criterio del profesorado + puntuación inicial en comprensión/ajuste (≤ 84) + puntuación inicial en <i>Cattell</i> (≤ 84)		
		PM	n	$r_{Y\lambda}$	PM	n	$r_{Y\lambda}$	PM	n	$r_{Y\lambda}$
ECO1	GC	+6.04*	36	.39	+13.98**	13	.76	+7.10	6	.51
	GE	+0.60	30	.07	+6.12*	11	.60	+6.73*	7	.73
ECOS	GC	+5.15**	34	.56	+8.06*	7	.75	+12.30*	4	.91
	GE	+7.93**	25	.74	+8.68**	12	.84	+5.79*	7	.78
APSL	GC	-2.22	56	-.15	+4.01	12	.31	+10.31	7	.67
	GE	+0.11	54	.00	+8.36*	11	.75	+7.60*	8	.73
Cattell	GC	+3.07	42	.25	-3.66	15	-.35	-3.00	11	-.39
	GE	+4.79**	57	.41	+7.09**	24	.69	+8.33**	15	.74

Nota: En los contrastes *pre-post* se utilizó la prueba *t* para muestras relacionadas; los criterios de análisis (b) y (c) para *APSL* contemplaron la puntuación inicial obtenida en la propia escala; PM = puntuación de mejora; * $p < .05$; ** $p < .01$.

Tabla 8

Resultados en GC y GE para la Puntuación de Mejora (Postest-Pretest) en las Distintas Pruebas, en Función de los Distintos Subgrupos Contemplados

Pruebas		DA			DIS			DEP			AE		
		PM	n	$r_{Y\lambda}$	PM	n	$r_{Y\lambda}$	PM	n	$r_{Y\lambda}$	PM	n	$r_{Y\lambda}$
ECO1	GC	+5.30	22	.35	+1.68	4	.20	+9.42	10	.52	+1.12	8	.12
	GE	-0.20	11	-.03	+1.28	7	.20	+0.93	12	.09	+0.23	19	.03
ECOS	GC	+5.24**	24	.53	+4.92*	10	.66	—	0	—	+5.15**	16	.75
	GE	+9.44**	17	.77	+5.54*	6	.80	+2.17	2	.51	+8.25**	17	.74
APSL	GC	-2.83	33	-.17	-2.30	16	-.30	+0.80	7	.04	-0.49	26	-.06
	GE	+2.04	29	.16	-4.89	13	-.48	+0.82	12	.09	+0.72	37	.06
Cattell	GC	+4.72	22	.34	+2.26	15	.26	-1.90	5	-.13	+2.69	23	.26
	GE	+2.77	30	.25	+10.64**	11	.74	+4.56	16	.41	+5.58**	36	.46

Nota: En los contrastes *pre-post* se utilizó la prueba *t* para muestras relacionadas; DA = dificultades de aprendizaje; DIS = discapacidad; DEP = privación socio-cultural; AE = recibe apoyo específico; PM = puntuación de mejora; * $p < .05$; ** $p < .01$.

Con relación a la prueba *ECOS* constatamos ganancias significativas en ambos grupos en función de los tres criterios de análisis. Los contrastes entre las PM mostraron diferencias significativas favorables al GE en función del primer criterio para la dimensión *Supervisión y autoevaluación del proceso de comprensión* ($t_{(2,57)} = -2,026, p < .047, r_{Y\lambda} = .26$). Asimismo, el GE obtuvo ganancias significativas en los análisis efectuados para los distintos subgrupos, a excepción del grupo DEP, aunque las diferencias promedio con relación al GC no llegaron a ser relevantes estadísticamente.

En cuanto a la escala *APSL*, los resultados muestran ganancias significativas en el GE en función del segundo y tercer criterio de análisis. Estas ganancias se produjeron en la puntuación global y especialmente en aquellas dimensiones relacionadas con el metaconocimiento estratégico sobre lectura (D8) y con el reconocimiento de la capacidad de generalizar estrategias (D9) («*Wilcoxon*»_(n=11) = 32.00, $p < .048$, en función del criterio (b) y «*Wilcoxon*»_(n=8) = 26.50, $p < .033$ para el criterio (c)). Por otra parte, hemos realizado para *APSL* contrastes en función de la etapa. Estos análisis nos permitieron comprobar que la dimensión D9 y aquella que expresaba conjuntamente las dimensiones D8 y D9, establecían diferencias significativas en E.S.O. (D9, $t_{(2,25)} = -2.848, p < .009, r_{Y\lambda} = .49$; D8 y 9, $t_{(2,25)} = -2.900, p < .008, r_{Y\lambda} = .50$). En ambos ca-

sos la medida del tamaño del efecto nos permite afirmar que en torno al 25% de la varianza en las puntuaciones postratamiento del GE ($r_{Y\lambda}^2$) sería explicada por la aplicación de *EDPL*. Por su parte, en el GC hubo ausencia de diferencias reseñables, tanto en primaria como en secundaria. Con relación a los contrastes entre las PM, en función del primer criterio hallamos diferencias significativas en secundaria para la dimensión D8 y 9 a favor del GE —en este contraste no se asumieron varianzas iguales ($t_{(2,58,62)} = -2.458, p < .017, r_{Y\lambda} = .29$)—. Para el segundo criterio, los contrastes de PM muestran diferencias notables favorables al GE, aunque no llegan a ser estadísticamente significativas ni para la puntuación global ni para las distintas dimensiones contempladas. Para el tercer criterio, en cambio, hallamos diferencias significativas favorables al GC para la dimensión D2 ($U_{(n=15)} = 3.00, p < .002$). Con relación a los distintos subgrupos, los contrastes efectuados para las distintas dimensiones contempladas mostraron diferencias significativas en el GE para el subgrupo DA en la dimensión compuesta (D8 y 9), con una ganancia postratamiento de 7.02 puntos ($t_{(2,28)} = -2.486, p < .02, r_{Y\lambda} = .43$). También para el subgrupo DA, hallamos diferencias significativas en la dimensión D1 en los contrastes de PM, con una diferencia promedio favorable al GE de 8.16 puntos ($t_{(2,60)} = -2.083, p < .042, r_{Y\lambda} = .26$). Para el resto de

subgrupos no hallamos diferencias promedio significativas.

Con relación a los resultados en Cattell, pudimos comprobar que para los distintos criterios adoptados las ganancias postratamiento del GE fueron muy significativas. Estas ganancias se produjeron tanto en primaria ($t_{(2,29)} = -2.245, p < .033, r_{Y\lambda} = .38$) como en secundaria ($t_{(2,26)} = -2.622, p < .014, r_{Y\lambda} = .46$). Los resultados positivos también se hicieron extensivos a los análisis efectuados en función de los distintos subgrupos contemplados, obteniendo ganancias muy significativas el subgrupo DIS ($t_{(2,10)} = -3.482, p < .006, r_{Y\lambda} = .74$) y el de AE ($t_{(2,35)} = 3.085, p < .004, r_{Y\lambda} = .46$). En cuanto a los contrastes entre las PM, éstos arrojaron diferencias significativas favorables al GE en función del segundo criterio ($t_{(2,37)} = 3.80, p < .001, r_{Y\lambda} = .52$) y también en función del tercero ($t_{(2,24)} = 3.71, p < .001, r_{Y\lambda} = .59$). Asimismo, los contrastes de PM muestran diferencias significativas para el subgrupo DIS favorables al GE. En estos contrastes hemos utilizado pruebas no paramétricas debido a las diferencias observadas en el pretest ($U_{(n=26)} = 41.00, p < .031$). Por su parte, los resultados para el GC en estos análisis se muestran inconsistentes.

Discusión

Los resultados obtenidos generan algunas cuestiones que merecen ser discutidas: ¿Cuáles han sido

los mecanismos de acción de *EDPL* que han permitido observar efectos positivos en algunos alumnos tras su aplicación? ¿Por qué se produjeron mejoras en rendimiento cognitivo con un dispositivo de ED de dominios específicos? Una peculiaridad de nuestro estudio la constituye el hecho de que el tratamiento experimental es, en sí mismo, un procedimiento de evaluación. El dispositivo *EDPL* se propone esencialmente obtener información válida con relación a la mejora del proceso de comprensión del alumnado con especiales dificultades; sin embargo, el proceso de *interacción mediada* que tiene lugar durante su aplicación, conlleva la puesta en práctica de una serie de *patrones metodológicos de evaluación dinámica* (PMED) que habrían facilitado la aparición de habilidades y estrategias de mayor nivel. En este sentido, los evaluadores extrajeron información acerca de aquellos PMED que se habrían mostrado eficaces (Navarro, 2008). Por otra parte, el análisis preferente de las habilidades metacognitivas ha conllevado igualmente una focalización de la actividad mediadora sobre procesos y estrategias de control y autorregulación del aprendizaje; actividad que podría haber facilitado igualmente la aparición, durante el proceso mismo de ED, de este tipo de habilidades en el alumnado y que parece haberse reflejado en los resultados, a través de las mejoras obtenidas en aquellas dimensiones de las pruebas-criterio más relacio-

nadas con la metacognición. Estos procesos serían, precisamente, los que parecen admitir en mayor medida una acción educativa eficaz para optimizar el funcionamiento cognitivo del alumnado con dificultades de aprendizaje o discapacidad intelectual (Carlson y Wiedl, 2000; García-Sánchez, 2002; Guterman, 2002; Navarro y Mora, 2009; Souvignier y Mokhlesgerami, 2006). En la introducción apuntábamos que algunas investigaciones en ED habían centrado esencialmente su intervención en estas habilidades y que los resultados obtenidos confirmaban su relevancia en la mejora del funcionamiento cognitivo de estudiantes con dificultades. En este sentido, Carlson y Wiedl (2000) encontraron que en la reestructuración de pruebas cognitivas, mediante la introducción de *retroalimentación detallada* y de *verbalización* antes, durante y después de la tarea, como factores clave en la mejora, eran precisamente aquellos estudiantes con dificultades metacognitivas, los que incrementaban significativamente su rendimiento. La mayor eficacia del modelo de evaluación propuesto por estos autores, residía de esta forma en la valoración dinámica de estas habilidades en los alumnos que presentaban déficit en estos procesos. La relevancia de la verbalización/mediación podría estar relacionada con una posible reestructuración de los procesos de pensamiento, en el sentido de optimizar los recursos atencionales, lo que podría afectar positivamente a la con-

sideración de las reglas que deben ser seguidas, las estrategias que deben ser empleadas o la información que debe ser recordada. Por otra parte, un segundo factor del proceso de verbalización que podría incidir notablemente en la mejora de la ejecución de las tareas, es que éste supone en la práctica el uso de un sistema de codificación adicional; en efecto, cuando se introduce la verbalización del proceso de aprendizaje, puede llegar a producirse una codificación doble y conjunta de la información, lo que redundará en un aprendizaje con mayores garantías. En nuestro caso, las mejoras postratamiento observadas en la prueba de inteligencia podrían ser explicadas desde esta hipótesis. De este modo, los estudiantes que obtuvieron ganancias posttest podrían haber utilizado inicialmente un sistema de codificación preferentemente perceptivo-visual, que se habría revelado insuficiente para dar cuenta con éxito de la tarea a realizar. Sólo posteriormente a la aplicación, habría tenido lugar una conjunción de dos sistemas de codificación; por una parte, el sistema perceptivo-visual, ya presente; y por otro, un sistema de codificación simbólico-lingüístico, fruto de la apropiación de estrategias de autorregulación verbal mediadas en el transcurso de la implementación del dispositivo. En este mismo sentido, a pesar de que la aplicación del dispositivo ha actuado esencialmente sobre procesos de lectura, y, por tanto, sobre un ámbito fundamentalmente

verbal; el trabajo desarrollado, basado especialmente en los procesos metacognitivos, y centrado por tanto más sobre estrategias generales de resolución de problemas (aunque aplicadas a un dominio específico), que sobre habilidades ligadas a una tarea concreta, podría explicar los cambios en la actuación con pruebas de contenido no verbal. Los estudiantes habrían aprendido estrategias para resolver problemas de diversa índole, porque la esencia de su aprendizaje estaría en la aplicación de elementos generales como, por ejemplo: (a) una mayor reflexión previa a la actuación; (b) el mantenimiento de la atención por más tiempo y con mayor calidad; (c) un mayor control y regulación sobre la actividad; o, (d) la revisión de su actividad una vez terminada. De esta forma, la aplicación de estas estrategias a la realización de una prueba como *Cattell*, habría contribuido de forma significativa a la explicación de las mejoras obtenidas en el postest.

Por otra parte, el diseño utilizado nos permitió controlar las principales fuentes de error: edad y rendimiento previo en inteligencia, ajuste personal-social y comprensión. Los niveles de equivalencia obtenidos contribuyeron a ejercer un control sobre los efectos principales de regresión a la media, que, de otro modo, hubieran puesto en cuestión las ganancias observadas en aquellos estudiantes que partían de puntuaciones más bajas. Sin embargo, a pesar de los resultados mencio-

nados con relación al rendimiento cognitivo, el dispositivo no ha generado efectos diferenciales en la comprensión. Debemos hacer mención, no obstante, a una limitación relativa al mayor número de alumnos que en el GE recibían AE debido a la presencia de mayores DL. Este hecho, a pesar de la equivalencia inicial en rendimiento lector, al menos en primaria, podría haber supuesto un hándicap importante con relación a efectuar una comparación válida inter-grupos para este aspecto. La razón de esto podría encontrarse en las mayores posibilidades de evolución favorable a corto/medio plazo en estudiantes que presentan un menor grado de dificultades (hecho que podría haber tenido lugar en el GC), en comparación con aquellos que manifiestan dificultades más específicas que requieren de una atención especial (Campione; Brown y Ferrara, 1987; Navarro y Mora, 2009). A pesar de esto, en relación a nuestra hipótesis, debemos concluir que *EDPL* habría originado mejoras significativas en aquellos estudiantes que presentaban mayores dificultades, tanto en funciones y dominios específicos, como en inteligencia general, al tiempo que nos habría otorgado información acerca de qué mecanismos podrían explicar las mejoras obtenidas. En este sentido, las implicaciones educativas que se derivan de estos resultados, especialmente en lo concerniente al aprendizaje de estrategias de autorregulación por parte de aquellos

sujetos con mayores dificultades, tanto en comprensión como en inteligencia, deben relacionarse necesariamente con la identificación de mecanismos de acción y de mediación, utilizados por los aplicadores durante la implementación del dispositivo, que se hubieran revelado claves, bien en relación a la obtención de información valiosa orientada a la intervención o bien como guía y apoyo que optimizó la actuación del sujeto (Grigorenko, 2009).

Los resultados estarían en consonancia con los obtenidos en investigaciones que mencionábamos en la *introducción* y que destacaban precisamente la relevancia de los procedimientos dinámicos de evaluación en el alumnado con mayores dificultades en su aprendizaje. De este modo, los datos obtenidos habrían configurado empíricamente la población que podría obtener mayores beneficios de la aplicación del dispositivo en el contexto escolar.

Referencias

- Aguilera, A., y García, I. (2004). Evaluar interviniendo, intervenir evaluando: una propuesta de actuación ante las dificultades del aprendizaje. *Apuntes de Psicología*, 22(3), 309-322.
- Beckmann, N., Beckmann, J. F., y Elliott, J. G. (2009). Self-confidence and performance goal orientation interactively predict performance in a reasoning test with accuracy feedback. *Learning and Individual Differences*, 19(2), 277-282.
- Botella, J., León, O., San Martín, R., y Barriopedro, M. I. (2001). *Análisis de datos en psicología I*. Madrid: Pirámide.
- Campione, J. C., y Brown, A. L. (1987). Linking dynamic testing with school achievement. En C. S. Lidz (Ed.), *Dynamic assessment: An international approach to evaluating learning potential* (pp. 82-115). New York: Guilford Press.
- Campione, J. C., Brown, A. L., y Ferrara, R. A. (1987). Retraso mental e inteligencia. En R. Sternberg, *Inteligencia humana, II* (pp. 599-729). Barcelona: Paidós.
- Carlson, J. S., y Wiedl, K. H. (2000). The validity of dynamic assessment. En C. S. Lidz y J. G. Elliot (Eds.), *Dynamic assessment: prevailing models and applications* (pp. 681-712). Greenwich, CT: Elsevier-JAI.
- Cattell, R. B., y Cattell, A. K. S. (1974/2001). *Test de factor «G» (Escalas 2 y 3)*. Madrid: TEA.
- Cioffi, G., y Carney, J. (1983). Dynamic assessment of reading disabilities. *Reading Teacher*, 36, 764-768.
- Coll, C., Barberà, E., y Onrubia, J. (2000). La atención a la diversidad en las prácticas de evaluación. *Infancia y Aprendizaje*, 90, 111-132.
- Compton, D., Fuchs, D., Fuchs, L., Bouton, B., Gilbert, J., Barquero, L., et al. (2010). Selecting at-risk first-grade readers for early intervention: eliminating false positives and ex-

- ploring the promise of a two-stage gated screening process. *Journal of Educational Psychology*, 102(2), 327-340.
- Dörfler, T., Golke, S., y Artelt, C. (2009). Dynamic assessment and its potential for the assessment of reading competence. *Studies in Educational Evaluation*, 35, 77-82.
- Elliott, J. (2003). Dynamic assessment in educational settings: realising potential. *Educational Review*, 55(1), 15-32.
- Fernández-Ballesteros, R., y Calero, M. D. (2000). The assessment of learning potential: the EPA instrument. En C. S. Lidz y J. G. Elliott (Eds.), *Dynamic assessment: prevailing models and applications* (pp. 293-323). Greenwich, CT: Elsevier-JAI.
- Feuerstein, R. (1996). La teoría de la modificabilidad estructural cognitiva. En S. Molina y M. Fandos (coords.), *Educación Cognitiva I* (pp. 31-75). Zaragoza: Mira Editores.
- García-Sánchez, J. N. (2002). *Aplicaciones de intervención psicopedagógica*. Madrid: Pirámide.
- Gersten, R., y Dimino, J. (2006). New directions in research RTI (response to intervention): rethinking special education for students with reading difficulties (yet again). *Reading Research Quarterly*, 41(1), 99-108.
- Graesser, A. C., Singer, M., y Trabasso, T. (1994). Constructing inferences during narrative test comprehension. *Psychological Review*, 101(3), 371-395.
- Grigorenko, E. (2009). Dynamic assessment and response to intervention. Two sides of one coin. *Journal of Learning Disabilities*, 42(2), 111-132.
- Grigorenko, E., Sternberg, R. J., y Ehrman, M. E. (2000). A theory-based approach to the measurement of foreign language learning ability: The CANAL-FT theory and test. *The Modern Language Journal*, 84, 390-405.
- Guterman, E. (2002). Toward dynamic assessment of reading: applying metacognitive awareness guidance to reading assessment tasks. *Journal of Research in Reading*, 25(3), 283-298.
- Guthke, J., y Beckmann, J. (2000). The learning test concept and its application in practice. En C. S. Lidz y J. G. Elliott (Eds.), *Dynamic assessment: prevailing models and applications* (pp. 17-69). Greenwich, CT: Elsevier-JAI.
- Hacker, D. (1998). Self-regulated comprehension during normal reading. In D. Hacher, J. Dunlosky y A. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 165-192). Hillsdale, N. J.: Lawrence Earlbaum Associates.
- Haywood, H. C., y Lidz, C. S. (2007). *Dynamic assessment in practice: clinical and educational applications*. New York: Cambridge University Press.
- Kaniel, S., Tzuril, D., Feuerstein, R., Ben-Shachar, N., y Eitan, T. (1991). Dynamic assessment, learning, and transfer abilities of Jewish Ethiopian immigrants to Israel. En R. Feuerstein, P. S. Klein y A. Tannenbaum (Eds.), *Mediated learning experience* (pp. 179-209). Londres: Freund.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95(2), 163-182.
- Kintsch, W., y Kintsch, E. (2005). Comprehension. En S. G. Paris y S. A. Stahl (Eds.), *Children's reading comprehension and assessment* (pp. 71-91). Mahwah, NJ: Lawrence Erlbaum.

- Kozulin, A. (2000). *Instrumentos psicológicos. La educación desde una perspectiva socio-cultural*. Barcelona: Paidós.
- Kozulin, A., y Garb, E. (2001). *Dynamic assessment of EFL text comprehension of at-risk students*. Paper presented at the 9th Conference of the European Association for Research on Learning and Instruction. Fribourg, Switzerland, 2001.
- Lidz, C. S. (2004). Comment on curriculum-based measurement: describing competence, enhancing outcomes, evaluating treatment effects, and identifying treatment nonresponders by Fuchs and Fuchs. *Journal of Cognitive Education and Psychology*, 4(1), 131-133.
- Lidz, C. S., y Elliott, J. G. (Eds.) (2000). *Dynamic assessment: prevailing models and Applications*. Greenwich, CT: Elsevier-JAI.
- Lukas, J. F., y Santiago, K. (2004). *Evaluación educativa*. Madrid: Alianza.
- Moreno, J., y Saldaña, D. (2005). Use of a computer-assisted program to improve metacognition in persons with severe intellectual disabilities. *Research in Developmental Disabilities*, 26, 341-357.
- Navarro, J. J. (2008). *Evaluación dinámica de procesos implicados en la lectura*. Tesis doctoral. Universidad de Sevilla.
- Navarro, J. J., y Mora, J. (2009). Metaconocimientos y comprensión de textos. *Investigación en la Escuela*, 69, 99-113.
- Navarro, J. J., y Mora, J. (2011). Analysis of the implementation of a dynamic assessment device of processes involved in reading with learning-disabled children. *Learning and Individual Differences*, 21(2), 168-175, doi: 10.1016/j.lindif.2010.11.008.
- Newman, D., Griffin, P., y Cole, M. (1989). *The construction zone*. New York: Cambridge University Press. (Trad. cast.: *La zona de construcción del conocimiento: trabajando por un cambio cognitivo en educación*. Madrid: Morata, 1998).
- Pardo, A., y San Martín, R. (1998). *Análisis de datos en psicología II*. Madrid: Pirámide.
- Rapp, D., y Broek, P. (2005). Dynamic text comprehension: an integrative view of reading. *Current Directions in Psychological Science*, 14(5), 276-279.
- Resing, W. C. M. (2000). Assessing the learning potential for inductive reasoning (LIR) in young children. En C. S. Lidz y J. G. Elliott (Eds.), *Dynamic assessment: prevailing models and applications* (pp. 224-262). Greenwich, CT: Elsevier-JAI.
- Resing, W., y Elliott, J. G. (2011). Dynamic testing with tangible electronics: Measuring children's change in strategy use with a series completion task. *British Journal of Educational Psychology*, 81, doi: 10.1348/2044-8279.002006.
- Resing, W., Tunteler, E., de Jong, F., y Bosma, T. (2009). Dynamic testing in indigenous and ethnic minority children. *Learning and Individual Differences*, 19(4), 445-450.
- Ruiz, U., y Camps, A. (2009). Investigar los géneros discursivos en el proceso educativo. *Revista de Psicodidáctica*, 14(2), 211-228.
- Sánchez, E. (1999). El lenguaje escrito y sus dificultades: una visión integradora. En A. Marchesi, C. Coll y J. Palacios (comps.), *Desarrollo psicológico y educación*, 3. *Trastornos del desarrollo y necesidades educativas especiales* (pp. 127-155). Madrid: Alianza.

- Saldaña, D., Mora, J., y Moreno, J. (2007). Evaluación dinámica en retraso mental severo: descripción de un procedimiento ecológico para la valoración de la metacognición. En J. N. García (coord.), *Dificultades del Desarrollo* (pp. 1-17). Madrid: Pirámide.
- Schlee, J. (1985). Förderdiagnostik - Eine bessere Konzeption? En R. S. Jäger, R. Horn y K. Ingenkamp (comps.), *Tests and trends* (pp. 82-208). Weinheim: Beltz.
- Souvignier, E., y Mokhlesgerami, J. (2006). Using self-regulation as a framework for implementing strategy instruction to foster reading comprehension. *Learning and Instruction*, 16(1), 57-71.
- Sternberg, R. J., y Grigorenko, E. L. (2002). *Dynamic Testing*. Cambridge: Cambridge University Press.
- [Evaluación dinámica. Naturaleza y mediación del potencial de aprendizaje. Barcelona: Paidós, 2003].
- Swanson, H. L., y Howard, C. B. (2005). Children with reading disabilities: does dynamic assessment help in the classification? *Learning Disability Quarterly*, 28(1), 17-34.
- Thurman, S. K., y McGrath, M. C. (2008). Environmentally based assessment practices: viable alternatives to standardized assessment for assessing emergent literacy skills in young children. *Reading & Writing Quarterly*, 24(1), 7-24.
- Vygotsky, L. S. (1934/1995). *Pensamiento y Lenguaje*. Barcelona: Paidós.
- Wells, G. (1999). Dialogic inquiry: Toward a socio-cultural practice and theory of education. Cambridge, UK: Cambridge University Press.

Juan José Navarro Hidalgo es Doctor en Psicología, Profesor Contratado en el departamento de Psicología Evolutiva y de la Educación de la Universidad de Sevilla. Su trabajo de investigación se centra en la evaluación dinámica de procesos implicados en la lectura, la metacognición y las dificultades de aprendizaje.

Joaquín Mora Roche es Doctor en Psicología, Profesor Titular en el departamento de Psicología Evolutiva y de la Educación de la Universidad de Sevilla. Dirige el grupo de investigación *Enriquecimiento cognitivo y necesidades educativas especiales*. Sus trabajos de investigación giran en torno a la discapacidad, la metacognición y el lenguaje. Es autor del programa de enriquecimiento cognitivo *Comprender y Transformar*.

Fecha de recepción: 11-1-11

Fecha de revisión: 23-3-11

Fecha de aceptación: 22-6-11

ANEXO I

Extracto de la sesión correspondiente al grupo de aplicación (B), CEIP Antonio Machado (profesora M.^a José Alés Hervás)

Transcripción correspondiente a la mediación realizada con C., un alumno de 5.º curso que presenta especiales dificultades en la comprensión global de textos. La actividad que sirve de base a esta mediación es la 22: «*Uso de la estructura de un texto*», perteneciente al bloque de procesos de integración textual. Esta actividad consistía básicamente en la lectura de una serie de textos por parte del alumnado, de los que posteriormente debía contar su contenido.

Dos de los textos que se presentaban eran de estructura narrativa, uno de ellos ordenado, respetando la estructura lógica, y el otro desordenado. Los otros dos textos eran de estructura expositiva y, al igual que en los narrativos, uno de ellos se presentaba ordenado y el otro no. Se trata de textos sencillos y con contenidos familiares; la dificultad de la tarea no debía estar en la complejidad de los contenidos de los textos. Se pretendía evaluar si los alumnos identificaban la estructura textual y si utilizaban elementos estructurales para acceder a la comprensión global de los textos. Asimismo, se quería valorar la estructuración del resumen del texto que realizaba el alumno, es decir, si realizaba un texto con sentido y estructura definidos.

Mediación efectuada	Análisis de los Patrones metodológicos de evaluación dinámica utilizados en la mediación
<p><i>Las ballenas (texto ordenado)</i></p> <p>Empieza a leer y vuelve atrás porque dice que no ha podido enterarse.</p> <p>C.: <i>¡He tenido que empezar tres veces por su culpa! (Su compañero, que no se centra en la actividad). Activa sus conocimientos previos cuando encuentra una palabra que vio en otro texto.</i></p> <p>C.: <i>¡Filtrar! Como la capa que rodea la tierra que filtra el sol.</i></p> <p>Pregunta las palabras que no entiende o desconoce. Controla la actividad, sabe que tiene que enterarse e introduce latencias para asegurarse. Vuelve a leer el texto, sin necesidad de mediación externa. Está supervisando la actividad. Explica el contenido del texto de forma coherente, con la secuencia adecuada. No hay mediación verbal externa por su parte.</p>	<p>La Evaluadora realiza anotaciones relativas a la puesta en práctica de procesos de autorregulación (estrategias de supervisión), procesos de ajuste personal-social y procesos de integración texto-conocimientos previos. El alumno pone en relación una de las palabras del nuevo texto con el contenido del texto anterior; sin embargo, no aparece recogida en la transcripción ninguna respuesta de refuerzo positivo o de confirmación de la respuesta o la actitud del sujeto por parte de la Evaluadora. Realiza anotaciones sobre la puesta en práctica de procesos de autorregulación (estrategias de supervisión). También realiza anotaciones sobre la cantidad y calidad de ayuda necesaria para acceder a la realización de la actividad</p> <p>*La Evaluadora realiza preguntas para facilitar</p>

Mediación efectuada	Análisis de los Patrones metodológicos de evaluación dinámica utilizados en la mediación
<p>Evaluadora: <i>¿Te ha gustado la lectura?</i> C.: <i>Sí</i> Evaluadora: <i>¿Te ha parecido fácil o difícil?</i> C.: <i>Fácil o regular.</i> Evaluadora: <i>Pero este texto ¿contaba bien las cosas?</i> C.: <i>Sí porque en el otro estaban las cosas mal, en el de Rabín. (Buscamos la lectura desordenada para recordarla).</i> Evaluadora: <i>Y desordenada qué pasaba, ¿era más fácil de entender?</i> C.: <i>No, más difícil, porque te lías.</i></p> <p><i>La respiración (texto desordenado)</i></p> <p>C.: <i>¿Por qué está el título aquí abajo?</i> Evaluadora: <i>¿Dónde tendría que estar?</i> C.: <i>Arriba.</i> Lee muy rápido. C.: <i>¡Ya! Pero lo voy a leer otra vez.</i> Durante la lectura hace comentarios acerca del contenido del texto. Evaluadora: <i>¿Has entendido todas las palabras del texto?</i> Pregunta varias palabras que no conocía, pero que espontáneamente no había preguntado. Cuando termina de leer: C.: <i>Este trozo no puede ser el primero porque dice: «en segundo...» Entonces tiene que venir otro primero.</i> Evaluadora: <i>Bien, ¿Y qué cuenta el texto?</i> C.: <i>No sé, es muy difícil.</i> Evaluadora: <i>¿De qué habla? ¿De ballenas?</i> C.: <i>No, de los pulmones.</i> Evaluadora: <i>¿Qué más?</i> C.: <i>Es muy difícil porque dice cosas de los pulmones, el oxígeno...</i> Leemos de nuevo párrafo a párrafo intentando comprender el significado global del texto. Le cuesta mucho trabajo, pero al final se queda con una leve idea. Sí detectó las incoherencias en la estructura del texto.</p>	<p>la autoevaluación del sujeto con relación a si le gustó la lectura, y a la facilidad o dificultad de su comprensión. *Pone en relación el texto que se está trabajando con el anterior, con el objetivo de provocar la evaluación del sujeto acerca de las características estructurales de ambos. *Facilita y acompaña la actividad del alumno, buscando el texto desordenado con el objeto de facilitar la comparación. *La Evaluadora establece apoyo gradual mediante la realización de preguntas que guían al sujeto hacia una conclusión.</p> <p>El alumno resalta una incoherencia en el texto relativa a la estructura, y la Evaluadora trata de *profundizar y cuestionar lo apuntado por el sujeto con una nueva pregunta. Realiza anotaciones sobre la conducta del sujeto relativa a la puesta en práctica de procesos de autorregulación (supervisión). *La Evaluadora realiza una cuestión que trata de provocar un proceso de reflexión y autoevaluación en el alumno con relación al conocimiento que tiene sobre el vocabulario del texto. *Pregunta al sujeto por el contenido del texto, *estableciendo apoyo gradual mediante nuevas preguntas ante la dificultad que supone para el sujeto relatar de qué va el texto. *Secuencia el contenido de la actividad para facilitar su realización. *La Evaluadora refuerza y confirma la respuesta del sujeto.</p>

NOTAS UTILIZADAS EN LA TRANSCRIPCIÓN:

Las anotaciones destacadas en negrita reflejan las reflexiones o comentarios de la propia evaluadora con relación al proceso de mediación o a la respuesta del alumno.

En la columna correspondiente al análisis efectuado, los asteriscos indican referencias a los patrones metodológicos de evaluación dinámica (PMED).