

# Conceptions of assessment when the teaching context and learner population matter: compulsory school versus non-compulsory adult education contexts

Ana Remesal · Gavin T. L. Brown

Received: 15 June 2014 / Revised: 26 September 2014 / Accepted: 20 November 2014 /  
Published online: 19 December 2014

© Instituto Superior de Psicologia Aplicada, Lisboa, Portugal and Springer Science+Business Media Dordrecht 2014

**Abstract** This paper presents results of a study on teachers' conceptions of assessment carried out on a sample of 493 teachers of Spanish as foreign language from all over the world. At the moment of data collection, the participants were members of an international online teacher community and were teaching at different professional contexts: basic compulsory school, and diverse extra-school teaching contexts, such as language academies and in-company education services. A self-report questionnaire was presented online to the teachers with 40 rating items capturing their conceptions of assessment and habitual assessment practices. Confirmatory factor analysis was undertaken to evaluate the theoretical model. Results reveal a new bifactor model of conceptions of assessment, which contrasts with the prevailing literature in the field. More specifically, results point to differences in teachers' conceptions of assessment depending on four aspects: how assessment affects the teaching process, the learning process, the certification of learning results, and the participants' accountability to different audiences, such as colleague teachers, families, and the general society. Our results further put forward that the learner population, whether children-adolescents or adults, influences teachers' conceptions of assessment. This should lead researchers and practitioners to rethink teachers' conceptions of assessment in a broader context than has been considered to date.

**Keywords** Conceptions of assessment · Compulsory school · Non-formal teaching · Teaching context · Adult learner

---

A. Remesal (✉)

Departament de Psicologia Evolutiva i de l'Educació, Facultat Psicologia, Universitat de Barcelona,  
Campus Mundet, Paseo del Valle Hebrón 171, E-08035 Barcelona, Spain  
e-mail: aremesal@ub.edu

G. T. L. Brown

Faculty of Education, The University of Auckland, Auckland 1142, New Zealand  
e-mail: gt.brown@auckland.ac.nz

## Introduction

The way teachers conceive of assessment purposes and practices is an important issue in education. Changes in assessment practices have proved to be a crucial instrument to promote deeper changes in teaching and learning practices (Barnes et al. 2000; Harlen 2005; Towndrow et al. 2010; Vandeyar and Killen 2003; Volante and Fazio 2007; Remesal 2006). However, the success of innovations may be either supported or resisted by the pre-existing beliefs of teachers who are expected to implement those changes (Fives and Buehl 2012). In this sense, we need to develop instruments that might be used by teacher educators, both in pre-service and in-service programs, to make teachers' conceptions of assessment explicit, thus accessible to reflection and change (DeLuca et al. 2013; Stiggins 1999), in order to develop assessment literacy (DeLuca and Klinger 2010). Teachers' conceptions of assessment have been an object of inquiry for over a decade within an intensive research program across different cultural contexts, among in-service teachers of primary or secondary schools (e.g. Delandshere and Jones 1999; Philippou and Christou 1997; Xu and Liu 2009; Brown 2004a; Remesal 2011; Coll and Remesal 2009), and also pre-service student-teachers (e.g. Ogan-Bekiroglu 2009; Wang et al. 2009; Brown and Remesal 2012).

In this study, we wanted to extend research to teachers in different contexts other than formal schooling. Foreign language teaching is, traditionally, a subject which transcends the formal, compulsory educational system. Hence, our target population turned to be foreign language teachers. Using a newly developed questionnaire (Remesal and Brown 2012), it was expected, in accordance with ecological rationality (Rieskamp and Reimer 2007), that conceptions of assessment would be influenced by teachers' working and policy context. First of all, we present the theoretical background of the inventory and the study. After that, we describe the instrument validation procedures with an international sample of 493 teachers of Spanish as a foreign language (SFL/*ELE*) working both in formal schooling and extra-schooling contexts (compulsory school, extra-school teaching, adult education and in-company training). We conclude with a discussion of the implications of this study.

## Theoretical background

The literature on teachers' conceptions of assessment has identified the tension between two main different purposes and uses of assessment in the compulsory educational system; that is, improvement or formative versus evaluative or summative (Coll and Remesal 2009). Several researchers refer to the same phenomena using different terms; for instance: testing versus assessment culture (Wolf et al. 1991), formative versus summative assessment (Black and Wiliam 1998), improvement versus accountability (Brown 2004a; Brown et al. 2011b) or educational regulation versus societal control (Perrenoud 2001; Remesal 2011). Although these terms are not necessarily 100 % synonymous, researchers agree that within educational systems, there are conflicts and tensions between agents, participants and stakeholders concerning the nature, purpose and effects of assessment.

Although the literature often presents these two poles as alternative or mutually exclusive views, we defend the dialectical nature of both options and the inherent necessity of both phenomena: the regulation of teaching and learning processes, on the one side, and the external control of and by various educational agents, on the other (Remesal 2011; Brown 2004a). However, there might be other factors that need to be considered. There is evidence of different interpretations of what 'improving teaching and learning by means of assessment' might mean in diverse cultural settings, as a result of different school traditions and legislation. For

example, in societies characterised strongly by public examinations (e.g. Hong Kong, China or Egypt), accountability is strongly correlated with improvement (e.g. Brown et al. 2011b; Brown et al. 2009; Gebril and Brown 2014). In other words, this systemic tension does have, in addition, a sociocultural component which affects assessment practices (Remesal 2007). Results of prior studies coincide in showing that conceptions of assessment tending towards formative practices are more frequently and strongly endorsed in primary schooling (e.g. Brown and Michaelides 2011; Brown et al. 2011a). In contrast, teachers in secondary education are more likely to regard assessment as an instrument of societal control or accountability rather than for improving teaching and learning processes in comparison with primary teachers (Remesal 2011).

Gaps identified in current research: our research question

Research to date has mainly focused on compulsory, formal systemic school settings at different compulsory levels (Liu 2008), but so we lack information on whether teachers' conceptions might follow similar patterns in other educational contexts. Only recent research tells us about differences between university lecturers and their students in the ways they conceive of assessment (Fletcher et al. 2012).

Hence, in this study, we looked at teachers working in different non-compulsory adult education teaching contexts, such as after-school language academies or in-company language courses, versus teachers in compulsory schooling or formal higher education, trying to find an answer to the question: Do teachers' conceptions about assessment in non-compulsory/non-formal adult education teaching contexts differ from those working in the compulsory school system? This research question is grounded on a twofold argument. First, there is the effect on teachers' conceptions of systemic policies and practices of assessment in compulsory education (Rasmussen and Friche 2011; Skedsmo 2010), whereas non-compulsory education is characterised by highly variegated policies and diversity of systems (e.g. private language schools, universities, corporate work-places, etc.). Second, basic schooling takes place during periods of childhood and adolescence. In contrast, extra-school and non-compulsory adult education contexts usually involve autonomous adults, with noticeable differences in terms of cognitive capacities, motivation, and self-regulation abilities (Gill 2011; Smoke 2013).

Theoretical model grounding the study

The theoretical model grounding the questionnaire rests on the definition of *conceptions* as the *organised subjective sum of individual beliefs*, which, in turn, are understood as *assumptions about objects and phenomena that people take as true* (often without intellectual contrast) (Green 1971; Pajares 1992). Beliefs composing an individual's conception of any particular topic differ from each other in being primary or central versus secondary or peripheral. Primary central beliefs are psychologically strong and resistant to change, whereas secondary and peripheral beliefs are more easily challenged. Furthermore, beliefs are arranged in clusters that may be held quite independently from each other. Green's model is particularly suitable for understanding the sometimes apparent inconsistency between beliefs and behaviour: while we express or espouse a certain belief, our current behaviour might be driven or enacted by another belief or set of beliefs which remain unspoken in the background.

The model was developed in earlier qualitative research and refers to two essential multi-faceted dimensions that shape how assessment is understood and evaluated (Remesal 2006; Remesal 2011). The first dimension refers to the focus of assessment: that is, how assessment *separately* may affect (a) teaching, (b) learning, (c) the certification of learning and (d) the

accountability of teaching. The second simultaneous dimension shaping teacher conceptions of assessment has to do with the overall control purpose of assessment: that is, assessment is either an *instrument for the regulation of educational processes* or an *instrument for societal control*. In the first option, assessment is a tool for reflection about teaching practices in the classroom, and/or as a tool for the improvement of students' learning. In the second option, assessment is limited to grading purposes and practices and is used to exercise control over teaching and learning; here, teachers are unlikely to show positive attitudes towards assessment as a tool for change.

In contrast with previous models (Brown 2004a), this model suggests that the beliefs teachers hold on the effects of assessment on teaching are indeed quite often separated from beliefs on how assessment affects learning. In other words, since teachers may conceive of teaching and learning as separated processes, some teachers might have a coherent conception of assessment, with aligned beliefs on assessment affecting both teaching *and* learning, whereas other teachers might as well have incoherent conceptions of assessment, with confronting beliefs on how assessment affects teaching and learning. More importantly, rather than relying solely on an inter-factor correlation between accountability and improvement to determine the relationship of improvement to accountability, this model takes a bifactor approach in which each item (whether it be about improved teaching or improved student learning) is also jointly predicted by either an evaluative or improvement-oriented perspective. This approach makes more explicit the effect of the external aspects of assessment on the internal educational functions.

As a matter of fact, the results of the original study from which the theoretical model was built (Remesal 2006) show that a coherent conception of assessment tending towards formative practices is more frequent in primary education, whereas teachers in secondary education are more likely to present incoherent conceptions, closer to regarding assessment as an instrument of societal control. However, this is not a rule of thumb, and the complexity of teachers' conceptions of assessment must be considered in relation with other aspects, such as previous education for teaching in primary and secondary school (which usually differs in every country), participation in professional development programs, teaching experiences and so forth (Liu 2008; Remesal 2011).

## Research method and design

This study, approved by the corresponding institutional ethical board, used a cross-sectional survey convenience sample of the target population of teachers of foreign language (Spanish in this case), with causal-correlational analysis of self-reported responses to a structured inventory.

## The questionnaire

We designed a 40-item self-report questionnaire with a positively packed agreement rating scale (i.e. two negative and four positive options) which has been found to be appropriate in conditions of social desirability (Brown 2004b). The questionnaire was presented to the members of an international online professional community (i.e. *ComunidadTodoole*, <http://www.todoole.net/>) of teachers of Spanish as foreign language (SFL). We had access to the members of this professional community in exchange for previous participation in an online seminar organised by the administrators of the community. Invitations to the community

members were sent three times in a 3-month period. Teachers participated voluntarily and for free. They had open online access to the questionnaire from their own computers and took about 30 min to complete it. Results of the study would be shared with the community once published, and the researchers' further participation in future seminars was also arranged in compensation of the participation in the study. The questionnaire is available at request to authors.

The items were aligned with the theoretical model that grounds the study, which was briefly presented in the previous section (Remesal 2006). Although the teacher conceptions of assessment inventory (Brown 2001–2003) already exists, a study with Spanish university students (Brown and Remesal 2012) demonstrated that the four-factor intercorrelated structure did not fit well with Spanish participants. Likewise, studies in Cyprus (Brown and Michaelides 2011) and Egypt (Gebril and Brown 2014) showed that the original statistical model developed in New Zealand did not apply outside that low-stakes assessment context. These supported the decision to test a different theoretical framework and a new instrument.

Each item was classified as belonging to one of four aspects focused by assessment (i.e. learning, teaching, accountability and certifying) as well as one of two purposes of assessment (i.e. formative regulation or societal control). The learning factor had to do with the motivational, feedback and regulatory effects assessment has on students, while the teaching factor had to do with the integration of assessment within curriculum and pedagogy. The accountability factor related to the use of assessment to evaluate and examine students, while certifying had to do with the use of assessments to establish standards and award certificates. Formative regulation speaks to the role assessment plays in informing and guiding teachers and learners to more effective learning outcomes, while societal control has to do with the role assessment plays in controlling teachers and students. Within each aspect of assessment, half the items were designed to reflect formative and half the controlling role of assessment.

## Participants

Altogether, 7500 teachers were invited to respond to the questionnaire. It had a response ratio of about 7 %; hence, the sample of 493 respondents has a margin of error of 4.72 %. The respondents were Spanish-as-foreign-language teachers from all over the world, working in different teaching contexts; that is, either compulsory secondary school or extra-school adult education language teachers. Basic demographic characteristics of the sample are presented in Table 1. The questionnaire was answered mainly by women (79 %), which is slightly over the average percentage of female teacher population in European countries in 2012<sup>1</sup> and the female teacher population in the UK specifically (DoE 2011), middle-aged (47 %) and those with >10 years teaching experience (48 %). Slightly more than half of the sample (55 %) taught adults in different contexts (either language private academies or in-company training), while the rest taught children or adolescents (either in compulsory school—primary to secondary—or in private language academies as an extra-after-school subject).

In this study, teaching experience, practice context and initial training were considered initial variables of comparison. The majority of the participants (60 %) had their initial training in the broad field of languages (e.g. translation, philology, linguistics); only one fourth in education (26 %) (e.g. pedagogy, didactics, educational psychology), whereas just 2 % of the sample had initial training in both fields. Finally, 12 % of the respondents could be considered 'intruders' in the profession since their initial training was apart from either language or education (e.g. graduates in History, Economics, or Engineering, etc.).

<sup>1</sup> <http://data.worldbank.org/indicator/SE.SEC.TCHR.FE.ZS/countries?display=default>

**Table 1** Demographic characteristics of sample

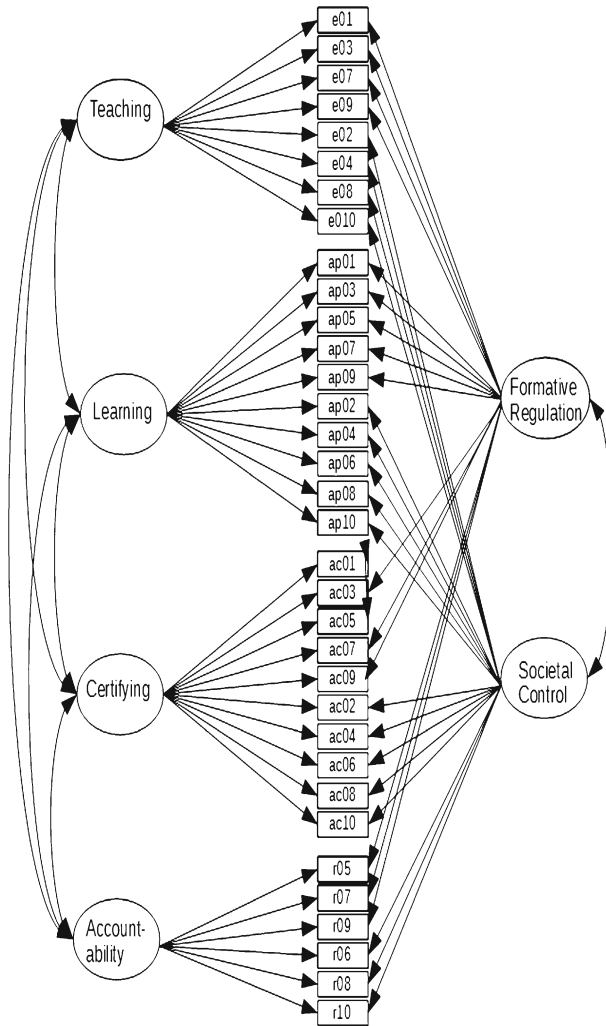
| Category                   | <i>N</i> | %  |
|----------------------------|----------|----|
| Age                        |          |    |
| <35                        | 160      | 32 |
| 35–50                      | 233      | 47 |
| >50                        | 100      | 20 |
| Sex                        |          |    |
| Male                       | 105      | 21 |
| Female                     | 288      | 79 |
| Teaching experience        |          |    |
| <2 years                   | 52       | 11 |
| 2–5 years                  | 96       | 19 |
| 5–10 years                 | 110      | 22 |
| >10 years                  | 235      | 48 |
| Initial training           |          |    |
| Linguistic-oriented        | 295      | 60 |
| Educational-oriented       | 128      | 26 |
| Linguistic and educational | 10       | 2  |
| Other                      | 59       | 12 |
| Student addressees         |          |    |
| Children-adolescents       | 223      | 45 |
| Adults                     | 270      | 55 |

## Analysis

Instead of resorting to exploratory factor analysis, the intended bifactor model was tested in confirmatory factor analysis. Bifactor models propose that each response is conditioned by two causes, while also modelling the effect of random error. While most bifactor models have a general common factor and independent group factors (Weekers 2009), the conceptual model proposed here consists of two dimensions (i.e. the focus of and the control purpose of assessment), each of which has two or more groups. This means that each item is modelled as being caused by both a focus and a control purpose factor (Fig. 1).

In line with current practice (Fan and Sivo 2007; Hu and Bentler 1999; Marsh et al. 2004), a multi-criteria approach for acceptable model fit was adopted; models were not rejected if gamma hat and comparative fit index (CFI)  $\geq 0.90$ , root mean square errors of approximation (RMSEA) and standardised root mean residuals (SRMR)  $\leq 0.08$  and  $\chi^2/df$  ratio were statistically non-significant ( $p < 0.05$ ). Models that met these criteria were not rejected. All analyses were carried out in AMOS (IBM 2011) using Pearson product moment correlations.

To determine the effect of teacher demographic characteristics upon conceptions of assessment, a factor mean score was created by bundling all the items predicted by the factor. Factor means were created instead of sums because of the different number of items in each factor. Multiple analyses of variance for the effect of key demographic variables on the eight interactive factor scores were conducted. Additionally, where MANOVA found statistically significant differences in mean scores for a factor, multi-group confirmatory factor analysis was carried out to determine whether the measurement model was statistically equivalent for that demographic variable. A nested, sequential approach first determines whether the model is configurally equivalent, then whether the regression weights are statistically equivalent, before



**Fig. 1** Bifactorial measurement model of 493 Spanish as a foreign language (*SFL*) teachers’ responses to inventory

testing the equivalence of item intercepts (Cheung and Rensvold 2002). Differences in the CFI>0.01 indicate that parameters are not equivalent. Lack of equivalence in the measurement model further reinforces the conclusion that the samples are drawn from different populations (Wu et al. 2007).

**Results**

Confirmatory factor analysis in which each item is predicted by both a focus and a purpose of assessment had acceptable fit ( $\chi^2=1040.39$ ,  $df=486$ ,  $[\chi^2/df=2.141$ ,  $p=0.14]$ ; CFI=0.81; gamma hat=0.94; RMSEA=0.048, 90 % CI=0.044–0.052; SRMR=0.052). While not every path was statistically significant, the combined weight of paths produced small to large



percentages of variance explained (min 2 % to max 47 %;  $M=25$  %,  $SD=12.5$  %) (see Table 2).

The mean scores for each of the factors (Table 3) showed the learning and accountability focus and the formative regulation purpose all received the highest means. However, inspection of the interaction means revealed a somewhat more instructive result. The formative regulation focus on learning and teaching had the highest means, while the societal control focus on the same two applications had the lowest means. All four formative regulation applications had higher means than the societal control applications, clearly indicating that the SFL teachers endorsed using assessment to guide pedagogy much more than for societal control reasons. It is noteworthy that the lowest mean (teaching SC) indicates that the conventional use of student assessment to evaluate the quality of teaching was more-or-less rejected.

Multiple analysis of variance of the effect of key demographic variables on conceptions used eight interactive factor mean scores (i.e. learning-FR, teaching-FR, accountability-FR, certifying-FR, learning-SC, teaching-SC, accountability-SC and certifying-SC). The predictor variables used were main effects for initial training and learner population (i.e. either children-adolescents or adults) and interaction between the two predictors. Statistically significant differences were found only for the main effect of working context (Wilks'  $\lambda=0.96$ ;  $F_{(8,478)}=2.71$ ,  $p=0.006$ ). Inspection of univariate differences for the eight factor scores showed that statistically significant differences existed for only three factors (i.e. accountability-SC,  $p=0.006$ ,  $d=0.26$ ; certifying-SC,  $p=0.031$ ,  $d=0.20$ ; and learning-SC,  $p<0.001$ ,  $d=0.38$ ). In all three cases, teachers in adult contexts gave lower scores, though the differences were small to moderate.

Multi-group confirmatory factor analysis found that after constraining the measurement weights to be the same for the two groups resulted in a difference of  $CFI=0.011$ . This supports the conclusion that the two groups responded to the inventory in different ways and that teachers of Spanish in these two different contexts are samples from two different populations.

## Discussion and conclusion

In this study, we wanted to learn about conceptions of assessment of teachers working in non-formal adult educational contexts versus those teaching in the formal and systemic school context, which led us to focus on teachers of Spanish as a foreign language. The results of this study show that, as proposed by the original study (Remesal 2006), teacher beliefs about assessment are organised in a complex model in which each individual evaluates the focus of assessment effects according to whether it has formative-regulatory or societal control uses. On the whole, this sample of SFL teachers conceived of assessment primarily in terms of its regulating effects on teaching and learning and hence was strongly positive towards formative assessment purposes. In other words, they agreed mostly that assessment exists to serve the improvement of learning and also teaching. In addition, we identified that those teachers working in non-compulsory adult contexts were less favourable towards the societal control purposes. In contrast, teachers in the formal compulsory schooling context showed a greater acknowledgement of assessment as a societal control tool by means of certifying achievement and showing accountability of teaching. This result might be explained by different elements that we address next.

First, teachers in formal schooling might have a greater awareness of their responsibility towards society and their own pupils or students through having to care for their basic development towards adulthood and responsible citizenship (Green et al. 2007; Johnson et al. 2008; Pope et al. 2009). In contrast, teachers in non-compulsory settings may regard this societal control by means of assessment as distant from their professional role (Vella



**Table 2** Item statistics by predictor factors

| Item | Spanish  | English  | Focus | $\beta$ | Control | $\beta$ | SMC  | <i>M</i> | SD    |
|------|--|--|-------|---------|---------|---------|------|----------|-------|
| ac01 | <i>Ante unos malos resultados de un grupo de alumnos, la mejor opción es bajar el nivel o incluso repetir la evaluación</i>                      | <i>In face of bad assessment results, the best option is to lower the standards or even to repeat the test.</i>                  | Cert  | 0.28    | SC      | 0.15    | 0.10 | 2.39     | 1.095 |
| ac02 | Una ‘bajada de nivel’ y hacer la ‘vista gorda’ nunca es la solución cuando nos encontramos con malos resultados de evaluación en una clase       | Lowering standards and blinking an eye are never the solution to face bad assessment results in a course.                        | Cert  | 0.24    | FR      | 0.23    | 0.11 | 4.06     | 1.153 |
| ac03 | <i>El avance de los alumnos siempre se debe valorar con el mismo rasero para todos, según los objetivos del curso planteados</i>                 | <i>Students’ learning progress must be always measured with the same scale for everyone, according to the course objectives.</i> | Cert  | 0.53    | SC      | 0.12    | 0.29 | 2.81     | 1.185 |
| ac04 | El punto de partida individual es la referencia imprescindible para valorar adecuadamente el avance de cada alumno                               | The individual starting point is an indispensable reference to properly evaluate the individual student’s progress.              | Cert  | 0.07    | FR      | 0.45    | 0.21 | 4.23     | 0.849 |
| ac05 | <i>Es preferible comunicar los resultados en forma estrictamente numérica, para evitar malentendidos</i>   | <i>It is always better to communicate assessment results in a numerical form, in order to avoid misunderstandings.</i>           | Cert  | 0.53    | SC      | 0.22    | 0.33 | 2.54     | 1.231 |
| ac06 | Las calificaciones numéricas (0–10, 6–1, 1–15,...), o categoriales básicas (A,B,C; aprobado, suspenso...) son, por lo general, poco informativas | Numerical grading or basic categories are, generally speaking, uninformative.  | Cert  | -0.26   | FR      | 0.18    | 0.10 | 3.48     | 1.123 |
| ac07 | <i>Para aprobar, el alumno debe alcanzar un dominio mínimo, indicado en el currículo oficial vigente</i>   | <i>In order to pass the course, the students must reach a minimal competence, indicated in the official curriculum.</i>          | Cert  | 0.56    | SC      | -0.18   | 0.35 | 3.89     | 0.964 |
| ac08 | Para aprobar, el alumno debe demostrar que ha avanzado bastante desde su punto de partida al inicio del curso                                    | In order to pass the course, the student must demonstrate a sufficient progress considering his own starting point.              | Cert  | 0.44    | FR      | 0.07    | 0.20 | 3.55     | 1.033 |

**Table 2** (continued)

| Item | Spanish  | English   | Focus | $\beta$ | Control | $\beta$ | SMC  | $M$  | SD    |
|------|--|---|-------|---------|---------|---------|------|------|-------|
| ac09 | <i>La comunicación de resultados siempre debe ser pública para que cada alumno se ubique en el grupo-clase</i> | <i>Communicating assessment results must always happen in public, so every student can locate himself in the class.</i>     | Cert  | 0.39    | SC      | 0.22    | 0.20 | 1.86 | 1.004 |
| ac10 | La comunicación de resultados debe ser siempre privada para evitar comparaciones                               | Communicating assessment results must always be private, to avoid comparisons among students.                               | Cert  | 0.03    | FR      | 0.19    | 0.04 | 3.55 | 1.214 |
| ap01 | <i>Los ‘exámenes sorpresa’ son un buen elemento motivador para los alumnos</i>                                 | <i>“Surprise exams” are a good motivational tool for students.</i>  | Lmg   | 0.32    | SC      | 0.21    | 0.14 | 2.15 | 1.103 |
| ap02 | El alumno siempre debe estar informado de la intención evaluativa del profesor                                 | The student must always be informed about the assessment intention of the teacher.  | Lmg   | 0.29    | FR      | 0.35    | 0.21 | 4.08 | 1.100 |
| ap03 | <i>Es imprescindible que los alumnos se concien del nivel que van alcanzando según el baremo establecido</i>   | <i>It is indispensable that the students get aware of their learning level according to the standards.</i>                  | Lmg   | 0.48    | SC      | -0.47   | 0.44 | 4.10 | 0.938 |
| ap04 | Es necesario que cada alumno sea consciente de su propio avance desde su propio punto de partida               | It is necessary that the students get aware of their own learning improvement, considering their particular starting point. | Lmg   | 0.04    | FR      | 0.61    | 0.37 | 4.64 | 0.608 |
| ap05 | <i>Evaluar el aprendizaje no les supone a los alumnos una ocasión de aprendizaje nueva</i>                     | <i>Learning assessment does not imply a new learning chance.</i>  | Lmg   | 0.20    | SC      | 0.49    | 0.28 | 1.99 | 1.149 |
| ap06 | Cada ocasión de evaluación es también una ocasión de posible nuevo aprendizaje                                 | Each assessment occasion is also a new opportunity for learning.  | Lmg   | -0.02   | FR      | 0.58    | 0.34 | 4.45 | 0.733 |
| ap07 | <i>Si no hubiera evaluación del aprendizaje sería imposible motivar a los alumnos para estudiar</i>            | <i>If there was not learning assessment, it would be impossible to motivate students to learn.</i>                          | Lmg   | 0.41    | SC      | -0.15   | 0.19 | 3.05 | 1.149 |
| ap08 | La evaluación vista como amenaza de castigo o promesa de premio no ayuda al aprendizaje                        | Assessment understood as a threat of punishment or promise of price does not help to learn.                                 | Lmg   | -0.03   | FR      | 0.12    | 0.02 | 3.98 | 1.209 |

**Table 2** (continued)

| Item | Spanish  | English  | Focus | $\beta$ | Control | $\beta$ | SMC  | $M$  | SD    |
|------|--|--|-------|---------|---------|---------|------|------|-------|
| ap09 | <i>Lo mejor para el alumno tras una evaluación negativa es hacer borrón y cuenta nueva</i>   | <i>The student's best option after bad assessment results is to make a clean slate.</i>  | Lmg   | 0.40    | SC      | 0.37    | 0.30 | 1.91 | 0.997 |
| ap10 | Cuando hay malos resultados de evaluación, el alumno debe reflexionar sobre sus errores  | If there are bad assessment results, the students must reflect on their mistakes.  | Lmg   | 0.24    | FR      | 0.29    | 0.14 | 4.26 | 0.937 |
| e01  | <i>La evaluación del aprendizaje y la enseñanza no se deben mezclar</i>  | <i>Learning assessment should not be mixed up with teaching.</i>   | Tchg  | 0.47    | SC      | 0.35    | 0.35 | 2.12 | 1.167 |
| e02  | La evaluación del aprendizaje y la enseñanza deben ser coherentes la una con la otra   | Learning assessment and teaching must be coherent.   | Tchg  | -0.20   | FR      | 0.65    | 0.45 | 4.47 | 0.761 |
| e03  | <i>La evaluación del aprendizaje supone una pérdida de tiempo y bastante estrés</i>  | <i>Assessment implies a loss of time and quite a stress.</i>   | Tchg  | 0.13    | SC      | 0.52    | 0.28 | 1.72 | 0.960 |
| e04  | La evaluación del aprendizaje de los alumnos es la brújula imprescindible de la enseñanza  | Learning assessment is the indispensable compass of teaching.  | Tchg  | 0.10    | FR      | 0.41    | 0.18 | 3.76 | 1.020 |
| e07  | <i>Las actividades de evaluación deben diferenciarse claramente de las actividades de aprendizaje</i>  | <i>Assessment activities and learning activities must be clearly differentiated.</i>   | Tchg  | 0.50    | SC      | 0.22    | 0.30 | 2.46 | 1.224 |
| e08  | Cualquier actividad de aprendizaje puede ser también actividad de evaluación   | Any learning activity may be as well used for the purpose of assessment.   | Tchg  | -0.17   | FR      | 0.68    | 0.11 | 4.08 | 1.015 |
| e09  | <i>La evaluación del aprendizaje no debe afectar a otras decisiones docentes (objetivos, contenidos, recursos, métodos...)</i>                                     | <i>Learning assessment must not affect other teaching decisions (objectives, contents, resources, methods...)</i>                                      | Tchg  | 0.50    | SC      | 0.39    | 0.40 | 2.22 | 1.212 |
| e10  | Cada vez que se evalúa el aprendizaje de los alumnos es también ocasión de revisar diversos aspectos de la enseñanza (objetivos, contenidos, recursos, métodos...) | Every time the students' learning is assessed, it is also the time to revise different aspects of teaching (objectives, contents, resources, methods). | Tchg  | -0.10   | FR      | 0.29    | 0.47 | 4.54 | 0.660 |

**Table 2** (continued)

| Item | Spanish  | English  | Focus | $\beta$ | Control | $\beta$ | SMC  | $M$  | SD    |
|------|--|--|-------|---------|---------|---------|------|------|-------|
| r05  | <i>La evaluación del dominio de ELE, pero particularmente la evaluación externa mediante los exámenes oficiales, es imprescindible para tener un control</i> | <i>Assessment of Spanish as Foreign Language, but particularly external assessment by means of official examinations, is indispensable to control.</i> | Acc   | 0.65    | SC      | 0.00    | 0.42 | 2.93 | 1.070 |
| r06  | Los exámenes oficiales de nivel de ELE ayudan a cada ciudadano a demostrar su propia competencia, por ejemplo ante futuros empleadores                       | Official SFL examinations help each citizen to demonstrate his own competence, for instance in front of likely employers.                              | Acc   | 0.44    | FR      | 0.22    | 0.24 | 3.74 | 1.019 |
| r07  | <i>El docente es el único responsable de elaborar o seleccionar actividades de evaluación y corregirlas posteriormente para los alumnos</i>                  | <i>The teacher is the only responsible of designing or selecting assessment activities and evaluating them afterwards for the students.</i>            | Acc   | 0.42    | SC      | 0.14    | 0.20 | 2.87 | 1.163 |
| r08  | Los alumnos son co-participes con el docente en todos los pasos de la evaluación, incluyendo la preparación de actividades y la corrección de resultados     | Students collaborate with the teacher in every step of the assessment process, including the preparation of activities and evaluating results.         | Acc   | -0.17   | FR      | 0.30    | 0.12 | 3.52 | 1.101 |
| r09  | <i>El currículo oficial determina los objetivos inexorables que cada docente debe cumplir</i>  | <i>The official curriculum determines the learning goals that every teacher must pursue.</i>   | Acc   | 0.67    | SC      | -0.08   | 0.46 | 3.12 | 1.057 |
| r10  | Cada docente debe ajustar su planificación de curso a los alumnos de su grupo  | The teacher must adjust his course plan to his group of students.  | Acc   | -0.13   | FR      | 0.47    | 0.23 | 4.32 | 0.801 |

*Italics text* = societal control

SMC squared multiple correlate showing % variance explained,  $\beta$  standardised regression weight, *Cert* certifying, *Lrng* learning, *Tchg* teaching, *Acc* accountability, *SC* societal control, *FR* formative regulation

1994). Particularly, in the field of foreign languages, as were the participants in this study, teaching focuses less upon basic knowledge or competences, but more upon individual opportunities for better employment or better accomplishment of personal motives and goals that guide the decision to enrol in a foreign language course (Salido 2006; Zúñiga 2009).

**Table 3** Mean scores for each factor

| Factor                     | <i>M</i> | <i>se</i> | SD   |
|----------------------------|----------|-----------|------|
| Focus of assessment        |          |           |      |
| Learning                   | 3.46     | 0.02      | 0.39 |
| Accountability             | 3.42     | 0.02      | 0.53 |
| Certifying                 | 3.23     | 0.02      | 0.44 |
| Teaching                   | 3.17     | 0.02      | 0.39 |
| Control purpose            |          |           |      |
| Formative regulation (FR)  | 4.03     | 0.012     | 0.40 |
| Societal control (SC)      | 2.61     | 0.02      | 0.51 |
| Focus *control interaction |          |           |      |
| Learning-FR                | 4.28     | 0.02      | 0.51 |
| Teaching-FR                | 4.21     | 0.03      | 0.56 |
| Accountability-FR          | 3.86     | 0.03      | 0.61 |
| Certifying-FR              | 3.77     | 0.03      | 0.56 |
| Accountability-SC          | 2.97     | 0.04      | 0.81 |
| Certifying-SC              | 2.70     | 0.03      | 0.66 |
| Learning-SC                | 2.64     | 0.03      | 0.57 |
| Teaching-SC                | 2.13     | 0.04      | 0.78 |

Secondly, we have to keep in mind the great distance between children or adolescents as learners in a compulsory context, and adult learners in a free context of self-development. The relationship of power between children or adolescents and the teachers shows a greater gap than that between the adult instructor and the adult learner freely choosing to involve in a learning experience (Lawler 2003; Pratt 1992). The adult learner has clearer motives for learning and has usually already developed strong learning strategies. In addition, the adult learner makes a certain economic investment in language courses, which in turn contributes to raised expectations for positive results and, hence, their expectation of greater formative interventions from the teacher (Knowles et al. 2012). Furthermore, a wider sympathy and empathy can develop between the adult teacher and the adult learner as equals in life phase (Nesbit 1998; Pratt 1992).

These results are coherent with the theoretical model which declares the twofold functional nature of assessment in the basic school system (Remesal 2011), which by default imposes both formative and non-formative (i.e. certification and accountability) purposes upon assessment. Within schooling contexts, teachers have to accept that assessment results are also used to evaluate students, monitor teacher effectiveness and judge school quality. In contrast, it appears that teachers working outside the basic school system might focus more on the classroom context and thus feel a lesser attachment or commitment to evaluative purposes.

It is also noteworthy that the strongly positive view towards formative regulation uses of assessment for teaching and learning is consistent with studies of compulsory education teachers in New Zealand (Brown 2004a; Brown and Michaelides 2011), Queensland (Brown et al. 2011a) and Hong Kong (Brown et al. 2009). Endorsing assessment as a mechanism for improved teaching and student learning does appear to be a universal attitude among societies in which teacher professionalism is an important facet of education. Like the previous study of pre-service teachers (Brown and Remesal 2012), this study found the least endorsement for the use of assessment to evaluate schools.

The current bifactor model is quite different to the multidimensional and hierarchical model of the widely used Teachers' Conceptions of Assessment inventory (Brown 2001–2003). Future research should investigate whether the models can be reconciled into a common framework. Further analysis of the non-equivalent aspects of the measurement model across working contexts is needed to see what further insights can be gained. Nonetheless, this study provides internal validation evidence for the questionnaire and advances our understanding of the contingencies in teachers' belief systems about assessment. Last but not least, it is remarkable that the literature on adult education mostly disregards the issue of assessment (Richey 1992; Knowles et al. 2012), even recent proposals of adults second language learning do not include the topic (Smoke 2013); merely contributions in the field of formal higher education and technical training tackle it (Arend 2009). It is hence time to include assessment into the research program of non-compulsory adult learning contexts. Our study is a first step in this direction, since we identified differences in teachers' conceptions that might be attributable to the educational context.

However, this study has limitations as well. Above all, there is an important limitation in the sample of participants; the huge variability of the sample, with no control over the particular institutional and cultural context, is certainly something to be addressed in future studies. In our study, we valued the voluntary participation and the diversity of ideas over sample control, but we recognise the methodological issue. Also, to continue this line of study, it will be necessary to focus on particular variables, like the sort of professional preparation that the teachers of foreign languages for adults receive on entry to teaching practice. The composition of our convenience sample did not allow us to make any deeper analysis into this; however, we found that there are three main backgrounds: linguistic, pedagogical and non-linguistic/non-pedagogical. It is our hypothesis that these different training backgrounds should generate different results on conceptions and practices of assessment. Finally, to make a comprehensive exploration of this new field, it would be necessary to pay attention to the other side of the road: the adult learners and their conceptions of assessment in relation to their own learning experience.

## References

- Arend, B. (2009). Using assessment to promote problem-solving and critical thinking skills in adults. In V. C. X. Wang (Ed.), *Assessing and evaluating adult learning in career and technical education*. Hershey: Information Science Reference.
- Barnes, M., Clarke, D. J., & Stephens, M. (2000). Assessment: the engine of systemic curricular reform? *Journal of Curricular Studies*, 32(5), 623–650.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74.
- Brown, G. T. L. (2001–2003). *Teachers' conceptions of assessment (TCoA) inventory* (Versions 1–3). Unpublished test. Auckland, NZ: University of Auckland.
- Brown, G. T. L. (2004a). Teachers' conceptions of assessment: Implications for policy and professional development. *Assessment in Education: Principles, Policy and Practice*, 11(3), 301–318.
- Brown, G. T. L. (2004b). Measuring attitude with positively packed self-report ratings: comparisons of agreement and frequency scales. *Psychological Reports*, 94(3), 1015–1024.
- Brown, G. T. L., & Michaelides, M. (2011). Ecological rationality in teachers' conceptions of assessment across samples from Cyprus and New Zealand. *European Journal of Psychology of Education*, 26, 319–337.
- Brown, G. T. L., & Remesal, A. (2012). Prospective teachers' conceptions of assessment: a cross-cultural comparison. *The Spanish Journal of Psychology*, 15(1), 75–89.

- Brown, G. T. L., Kennedy, K. J., Fok, P. K., Chan, J. K. S., & Yu, W. M. (2009). Assessment for student improvement: understanding Hong Kong teachers' conceptions and practices of assessment. *Assessment in Education: Principles, Policy & Practice*, 16(3), 347–363.
- Brown, G. T. L., Lake, R., & Matters, G. (2011a). Queensland teachers' conceptions of assessment: the impact of policy priorities on teacher attitudes. *Teaching and Teacher Education*, 27(1), 210–220.
- Brown, G. T. L., Hui, S. K. F., Yu, F. W. M., & Kennedy, K. J. (2011b). Teachers' conceptions of assessment in Chinese contexts: a tripartite model of accountability, improvement, and irrelevance. *International Journal of Educational Research*, 50(5–6), 307–320.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233–255.
- Coll, C., & Remesal, A. (2009). Concepciones del profesorado de matemáticas acerca de las funciones de la evaluación del aprendizaje en la educación obligatoria. *Infancia y Aprendizaje*, 32(3), 391–404.
- Delandshere, G., & Jones, J. H. (1999). Elementary teachers' beliefs about assessment in mathematics: a case of assessment paralysis. *Journal of Curriculum and Supervision*, 14(3), 216–240.
- DeLuca, C., & Klingner, D. A. (2010). Assessment literacy development: identifying gaps in teachers candidates' learning. *Assessment in Education: Principles, Policy and Practice*, 17(4), 419–438.
- DeLuca, C., Chavez, T., & Cao, C. (2013). Establishing a foundation for valid teacher judgement on student learning: the role of pre-service assessment education. *Assessment in Education: Principles, Policy and Practice*, 20(1), 107–126.
- DoE. (2011). *A profile of teachers in England from the 2010 School Workforce Census (Research Report DFE-RR151)*. London: Department of Education, Education Standards Analysis & Research Division.
- Fan, X., & Sivo, S. A. (2007). Sensitivity of fit indices to model misspecification and model types. *Multivariate Behavioral Research*, 42(3), 509–529.
- Fives, H., & Buehl, M. M. (2012). Spring cleaning for the “messy” construct of teachers' beliefs: what are they? Which have been examined? What can they tell us? In K. R. Harris, S. Graham, & T. Urdan (Eds.), *APA Educational Psychology Handbook: Individual Differences and Cultural and Contextual Factors* (Vol. 2, pp. 471–499). Washington, DC: APA.
- Fletcher, R. B., Meyer, L., Anderson, H., Johnston, P., & Rees, M. (2012). Faculty and students conceptions of assessment in higher education. *Higher Education*, 64(1), 119–133. doi:10.1007/s10734-011-9484-1.
- Gebriel, A., & Brown, G. T. L. (2014). The effect of high-stakes examination systems on teacher beliefs: Egyptian Teachers' Conceptions of Assessment. *Assessment in Education: Principles, Policy and Practice*, 21(1), 16–33. doi:10.1080/0969594X.2013.831030.
- Gill, D. (2011). Meeting differing learning styles of non-traditional students in the second language classroom. *Journal of College Teaching & Learning (TLC)*, 2(8), 1–8.
- Green, T. F. (1971). *The activities of teaching*. New York: McGraw-Hill.
- Green, S. K., Johnson, R. L., Kim, D. H., & Pope, N. S. (2007). Ethics in classroom assessment practices: issues and attitudes. *Teaching and Teacher Education*, 23(7), 999–1011.
- Harlen, W. (2005). Teachers' summative practices and assessment for learning—tensions and synergies. *The Curriculum Journal*, 16(2), 207–223.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55.
- IBM (2011). *Amos [computer program] (Version 20, Build 817)*. Meadville, PA: Amos Development Corporation.
- Johnson, R. L., Green, S. K., Kim, D. H., & Pope, N. S. (2008). Educational about ethical practices in leaders' perceptions student evaluation. *American Journal of Evaluation*, 29(4), 520–530.
- Knowles, M. S., Holton, E. F., & Swanson, R. A. (2012). *The adult learner*. Oxford: Routledge.
- Lawler, P. A. (2003). Teachers as adult learners: a new perspective. *New Directions for Adult and Continuing Education*, 2003(98), 15–22.
- Liu, X. (2008). Measuring teachers' perceptions of grading practices: does school level make a difference? *Proceedings of the Northeastern Educational Research Association (NERA) Annual Conference*, Rocky Hill, Connecticut. [http://digitalcommons.uconn.edu/nera\\_2008/4/](http://digitalcommons.uconn.edu/nera_2008/4/)
- Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, 11(3), 320–341.
- Nesbit, T. (1998). Teaching in adult education: opening the black box. *Adult Education Quarterly*, 48(3), 157–170. doi:10.1177/074171369804800303.
- Ogan-Bekiroglu, F. (2009). Assessing assessment: examination of pre-service physics teachers' attitudes towards assessment and factors affecting their attitudes. *International Journal of Science Education*, 31(1), 1–39.



- Pajares, M. F. (1992). Teachers' beliefs and educational research: cleaning up a messy construct. *Review of Educational Research*, 62(3), 307–332. doi:10.3102/00346543062003307.
- Perrenoud, P. (2001). Évaluation formative et évaluation certificative: postures contradictoires ou complémentaires? *Formation Professionnelle Suisse*, 4, 25–33.
- Philippou, G., & Christou, C. (1997). Cypriot and Greek primary teachers' conceptions about mathematical assessment. *Educational Research & Evaluation*, 3(2), 140–159.
- Pope, N., Green, S. K., Johnson, R. L., & Mitchell, M. (2009). Examining ethical dilemmas in classroom assessment. *Teaching and Teacher Education*, 25, 778–782.
- Pratt, D. D. (1992). Conceptions of teaching. *Adult Education Quarterly*, 42(4), 203–220. doi:10.1177/074171369204200401.
- Rasmussen, A., & Friche, N. (2011). Roles of assessment in secondary education: participant perspectives. *Educational Assessment, Evaluation and Accountability*, 23(2), 113–129.
- Remesal, A. (2006). Los problemas en la evaluación del aprendizaje matemático en la educación obligatoria: perspectiva de profesores y alumnos. Unpublished Doctoral Thesis. Universitat de Barcelona
- Remesal, A. (2007). Educational reform and primary and secondary teachers' conceptions of assessment: the Spanish instance. Building upon Black and Wiliam (2005). *The Curriculum Journal*, 18(1), 27–38.
- Remesal, A. (2011). Primary and secondary conceptions of assessment: a qualitative study. *Journal of Teaching and Teaching Education*, 27(2), 472–482.
- Remesal, A., & Brown, G.T.L. (2012). Spanish as Foreign Language Teachers' conceptions of assessment. Preliminary results from an internet inquiry. Paper presented at EARLI-SIG Assessment: Linking multiple perspectives of assessment. Brussels.
- Richey, R. (1992). *Designing instruction for the adult learner: systematic training theory and practice*. London: Kogan Page.
- Rieskamp, J., & Reimer, T. (2007). Ecological rationality. In R. F. Baumeister & K. D. Vohs (Eds.), *Encyclopedia of Social Psychology* (pp. 273–275). Thousand Oaks: Sage.
- Salido, M. H. (2006). ELE con fines humanitarios: nuevos espacios en la enseñanza de la Pragmática. In *La competencia pragmática y la enseñanza del español como lengua extranjera*. (pp. 379–388). Servicio de Publicaciones.
- Skedsmo, G. (2010). Formulation and realisation of evaluation policy: inconsistencies and problematic issues. *Educational Assessment, Evaluation and Accountability*, 23(1), 5–20.
- Smoke, T. (Ed.). (2013). *Adult ESL. Politics, pedagogy, and participation in classroom and community programs*. Oxford: Routledge.
- Stiggins, R. (1999). Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practice*, 18(1), 23–27.
- Towndrow, P. A., Tan, A.-L., Yung, B. H., & Cohen, L. (2010). Science teachers' professional development and changes in science assessment practices: what are the issues? *Research in Science Education*, 40, 117–132. doi:10.1007/s11165-008-9103-z.
- Vandeyar, S. & Killen, R. (2003). Has curriculum in South Africa really changed assessment practices, and what promise does the revised National Curriculum Statement hold? *Perspectives in Education*, 21(1), 119–134. <http://hdl.handle.net/2263/5044>
- Vella, J. (1994). *Learning to listen, learning to teach. The power of dialogue in educating adults*. San Francisco: Jossey-Bass.
- Volante, L., & Fazio, J. (2007). Exploring teacher candidates' assessment literacy: implications for teacher education reform and professional development. *Canadian Journal of Education / Revue Canadienne de l'Éducation*, 30(3), 749–770.
- Wang, J. R., Kao, H. L., & Lin, S. W. (2009). Preservice teachers' initial conceptions about assessment of science learning: the coherence with their views of learning science. *Teaching and Teacher Education*, 26(3), 522–529.
- Weekers, A. M. (2009). *Modeling typical performance measures*. (Unpublished doctoral dissertation), University of Twente, Enschede, NL.
- Wolf, D., Bixby, J., Glenn, J., & Gardner, H. (1991). To use their minds well: Investigating new forms of student assessment. *Review of research in education*, 17, 31–74.
- Wu, A. D., Li, Z., & Zumbo, B. D. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. *Practical Assessment, Research & Evaluation*, 12(3), Available online: <http://pareonline.net/getvn.asp?v=12&n=13>.
- Xu, Y., & Liu, Y. (2009). Teacher assessment knowledge and practice: a narrative inquiry of a Chinese college EFL teacher's experience. *TESOL Quarterly*, 43(3), 493–513.

Zúñiga, M. V. (2009). El aprendizaje de ELE para niños y adolescentes inmigrantes. *MarcoELE: Revista de didáctica*, 8, 12.

**Dr. Ana Remesal.** Departament de Psicologia Evolutiva i de l'Educació, Facultat Psicologia, Universitat de Barcelona, Campus Mundet, Paseo del Valle Hebrón 171, E-08035. Fax +34934021368, Email: aremesal@ub.edu

*Current themes of research:*

Teacher education. Assessment. Conceptions.

*Most relevant publications in the field of Psychology of Education:*

- Lafuente, M.; Remesal, A.; Álvarez Valdivia, I. (2013). Assisting Learning in e-Assessment: A Closer Look at Educational Supports. *Assessment and Evaluation in Higher Education*. doi:10.1080/02602938.2013.848835
- Brown, G.; Remesal, A. (2012). Prospective teachers' conceptions of assessment: a cross-cultural comparison. *The Spanish Journal of Psychology*, vol.15 (1).
- Remesal, A. (2011). Primary and secondary teachers' conceptions of assessment: A qualitative study. *Journal of Teaching and Teacher Education*, 27(2), 472–482.
- Coll, C.; Remesal, A. (2009). Concepciones del profesorado de matemáticas acerca de las funciones de la evaluación del aprendizaje en la educación obligatoria. *Infancia y Aprendizaje*, 32(3), 391–404.
- Remesal, A. (2007). Educational reform and primary and secondary teachers' conceptions of assessment: the Spanish instance. Building upon Black and Wiliam (2005). *The Curriculum Journal*, 18(1), 27–38.

**Dr. Gavin T. L. Brown.** H - Block, Epsom, Level 4, Room 6EH409; Epsom Campus, 74 Epsom Ave.; Auckland 1023, Epsom, New Zealand. Email: gt.brown@auckland.ac.nz

*Current themes of research:*

Self-assessment. Psychology of assessment. Assessment

*Most relevant publications in the field of Psychology of Education:*

- Panadero, E., Brown, G. T. L., & Courtney, M. G. (2014). Teachers' reasons for using self-assessment: A survey self-report of Spanish teachers. *Assessment in Education: Principles, Policy and Practice*. doi:10.1080/0969594X.2014.919247
- Brown, G. T. L., Irving, S. E., & Keegan, P. J. (2014). An introduction to educational assessment, measurement, and evaluation: Improving the quality of teacher-based assessment (3rd ed). Auckland, NZ: Dunmore Publishing. ISBN: 9781927212097 URL: <http://tinyurl.com/l4xtyzb>
- Brown, G. T. L., & Harris, L. R. (2014). The future of self-assessment in classroom practice: Reframing self-assessment as a core competency. *Frontline Learning Research*, 3, 22–30. doi: 10.14786/flr.v2i1.24
- Gebriel, A., & Brown, G. T. L. (2014). The effect of high-stakes examination systems on teacher beliefs: Egyptian Teachers' Conceptions of Assessment. *Assessment in Education: Principles, Policy and Practice*, 21(1), 16–33. doi: 10.1080/0969594X.2013.831030
- Chen, J., & Brown, G. T. L. (2013). High-stakes examination preparation that controls teaching: Chinese prospective teachers' conceptions of excellent teaching and assessment. *Journal of Education for Teaching*, 39(5), 541–556. doi: 10.1080/02607476.2013.836338
- Brown, G. T. L. (2013). Assessing Assessment for Learning: Reconsidering the policy and practice. In M. East & S. May (Eds.), *Making a Difference in Education and Social Policy* (121–137). Auckland, NZ: Pearson.