



Portuguese university students' conceptions of assessment: taking responsibility for achievement

Maria Assunção Flores¹ · Gavin Brown² · Diana Pereira¹ · Clara Coutinho¹ · Patrícia Santos¹ · Cláudia Pinheiro¹

Published online: 14 June 2019
© Springer Nature B.V. 2019

Abstract

After 20 years of the Bologna Declaration, Portuguese universities claim to be implementing student-centred and involved assessment practices. Student conceptions of assessment matter when innovations and reforms in assessment practices are being implemented. This study is part of a larger research project entitled “Assessment in higher education: the potential of alternative methods” funded by the Portuguese Foundation for Science and Technology (Government Funding Agency) (PTDC/MHCCED/2703/2014). This paper surveys a large sample ($N = 5549$) of Portuguese students in five public universities with a Portuguese version of the Students Conceptions of Assessment (SCoA-VI) inventory, previously validated in Brazil. Confirmatory factor analysis recovered the eight SCoA factors reported in the Brazilian context. Differences in mean score for the eight factors were trivial for institutional and student factors. Overall, students agreed that assessment for student improvement was not ignored. Scale inter-correlations revealed interesting inverse relations between improvement and irrelevance functions.

Keywords Conceptions of assessment · Higher education · Students' conceptions · Improvement · Accountability

Introduction

Assessment in higher education summatively evaluates and credentials student learning and, ideally, formatively supports improved learning. The influence of assessment on students' learning and performance is profound (Biggs 2003; Boud and Falchikov 2007; Brown and Knight 1994; Brown et al. 1997; Gulikers et al. 2008). Learning for a summative test or

✉ Maria Assunção Flores
aflores@ie.uminho.pt

¹ Research Centre on Child Studies (CIEC), Institute of Education, University of Minho, Braga, Portugal

² Faculty of Education and Social Work, University of Auckland, Auckland, New Zealand

qualification, however, is often seen as less than ideal because of its lack of emphasis on learning (Black and Wiliam 1998; Dochy and McDowell 1997; Myers and Myers 2015; Webber 2012). Higher education institutions are globally seeking to prioritise positive effects and benefits for learning through reformed assessment processes (e.g. rubrics, transparency, feedback, authenticity) and the use of innovative assessment practices (e.g. portfolios, self and peer assessment, simulation) (Struyven and Devesa 2016; Struyven et al. 2005).

The conceptions students have of those functions, including the value they give to each function, matter to the quality of their learning outcomes (McMillan 2016). Awareness of student beliefs especially matters when institutional policies and practices are reformed, partly because students tend to resist innovations in the mechanisms used to judge, evaluate, or certify their achievements (Struyven and Devesa 2016). As universities seek to innovate in their assessment regimes, it is important that students believe such assessments contribute to improved outcomes. Rejection of evaluative practices as irrelevant, inaccurate, or invalid would probably undermine the constructive goal of using assessment for learning. Hence, this paper extends the growing body of research into student conceptions of assessment by surveying a large sample of Portuguese university students with a recognised research tool (i.e. Student Conceptions of Assessment (SCoA) inventory; Brown 2008). Beyond validating that tool in a new context, the study explicitly tests the impact of relevant student and institutional factors upon the stability of results and differences in mean scores.

Students' conceptions of assessment in higher education

According to Segers and Tillema (2011, p. 53), “studying conceptions of assessment is of utmost relevance at a time when innovation of assessment practices is on the educational agenda”. Students' conceptions of assessment have been shown to relate significantly to academic performance, partly because of their contribution to self-regulation of learning (Brown 2011). Endorsement of assessment as a tool for improvement appears to be a constructive self-regulating belief leading to higher academic performance because self-regulation requires reflection upon achieved performance to identify learning priorities and successes. Gulikers et al. (2008) also found that students who perceived assessment as more authentic studied harder and developed more professional skills. In contrast, endorsement that assessment is irrelevant and ignorable leads to maladaptive responses (e.g. attribution to external locus of control) and rejection of the legitimacy of either the evaluative process, results, or feedback.

Students' conceptions and beliefs about assessment may influence their motivation and the process of self-regulated learning (Pereira 2016; Zimmerman 2008) because of the overlap between such ideas and self-regulation of learning (Brown 2011). According to Novak and Johnson (2012), emotions are directly linked to cognition and there is a strong correlation between affect and learning. Those conceptions lead to different reactions and feelings before, during, and after assessment (Boud 1995; McMillan 2016; Race 1995). Research also shows that students experienced both positive and negative emotions in higher education contexts (Novak and Johnson 2012). In addition, recent studies on assessment feedback reveal that emotional reactions may determine how students will perform regarding the feedback received (Pitt and Norton 2017; Ryan and Henderson 2018). In contrast, an intensive longitudinal study of student emotional responses through a summative assessment found that only when scores were known (i.e. feedback was received) did student emotions have systematic relations with performance (Peterson et al. 2015).

Much of the empirical work on conceptions of assessment has focused on compulsory education students and how these conceptions affect students' study behaviours and outcomes (Brown 2011, 2013; Brown and Harris 2012; Brown and Hirschfeld 2008; Brown et al. 2009b; Chen and Brown 2018; Otunuku et al. 2013; Peterson and Irving 2008; Solomonidou and Michaelides 2017). Segers and Tillema (2011) found gaps in research regarding how the conceptions of assessment influence students' learning and teachers' approaches to teaching in different assessment systems. Furthermore, according to Solomonidou and Michaelides (2017), although there is research evidence on students' conceptions regarding school improvement and evaluation of teaching, students' conceptions of their own assessment continue to be an under-researched area. While some research with higher education students does exist (Brown 2013; Brown et al. 2014; Brown and Wang 2013, 2016; Fletcher et al. 2012; Landim et al. 2015; Matos et al. 2009, 2013; Wang and Brown 2014), much less has been conducted in Portugal.

In the Portuguese context, research on assessment in higher education is still scarce. Flores and colleagues (Flores et al. 2015; Pereira et al. 2017a, b) found that written tests were the most used assessment methods and were associated with surface approaches to learning. Furthermore, the ideas which students associated most with assessment were linked to the kinds of assessment methods used by the teachers and were dependent on the programmes in which students are enrolled (Pereira et al. 2017a, b). Students perceived assessment as more effective and fairer when learner-centred assessment methods were used rather than only traditional assessment methods with implications for student learning processes and connections to their professional world (Flores et al. 2015; Pereira et al. 2017a). Learner-centred methods include, for instance, portfolios and project-based assessments, and "activities such as multiple drafts of written work that provides progressive feedback, oral presentations, student evaluations of each other's work and group and team projects that require interactions" (Webber (2012 p. 201).

In general, students associated most frequently neutral ideas (such as tests or examinations and grades) with assessment but also positive ideas (e.g. learning), while negative ideas (e.g. unfairness and fear) were least frequent. However, the idea of conflict was identified by undergraduate students who were assessed most frequently through so-called learner-centred methods of assessment such as project-based work and portfolios (Flores et al. 2015). However, differences among different programmes were found; for example, students enrolled in the programmes of Social Sciences and Humanities associated more positive ideas to assessment than students from other programmes (Flores et al. 2015).

Similar results were found by Fernandes and colleagues who conducted comparative studies of assessment in higher education in Portugal and Brazil (Fernandes 2014, 2015; Barreira et al. 2017). These studies revealed that university students perceived assessment as being essentially summative. However, traditional assessment methods (exams/tests) were not seen by the students as the best way of knowing what students learn and what they are capable of doing. In addition, students considered that, although assessment practices were continuous throughout the semester, there was little involvement in student-involved self- and peer assessment (Fernandes 2015; Barreira et al. 2017). Although the assessment system in Portugal is consistent with research-based recommendations and recognises the fact that formative assessment promotes students' educational success, in many Portuguese educational settings, assessment is more oriented towards grading and ranking pupils' achievements rather than on helping them to learn (Fernandes 2009).

The same emphasis on traditional methods of assessment is reported in Brazil (De Oliveira and Flores 2017). Undergraduate students state that written tests are the most common assessment methods and that grades usually have no feedback from the instructors. In a similar vein, Pereira et al. (2016b) showed that students perceive feedback as more relevant and effective when students are assessed through learner-centred methods rather than by traditional ones. In another study, comparing students' perceptions about assessment in Portugal and Sweden found similar views although student modes of assessment (i.e. self- and peer assessment) were less used by Swedish students (Pereira et al. 2017b). Likewise, although there is a large variation in assessment practices between universities and disciplines within neighbouring Spain, the use of traditional examination practices was widespread (Panadero et al. 2018). Thus, it could be expected that specific context factors could impinge upon student conceptions in Portugal.

A review of Brazilian research into higher education student perspectives on assessment (Matos et al. 2009) showed that two dimensions (i.e. formality of assessment method and locus of control) were able to summarise the variation in student thinking. Ultimately, assessment was perceived as formal test-like practices under the control of teachers, rather than students. Even peer- and self-assessment practices were directed by teachers, not students themselves. A large scale survey of public and private university students in Minas Gerais state in Brazil (Matos et al. 2012) found that students had a wide range of conceptions but their dominant conception was a negative emotional reaction (irrelevance) with respect to student accountability. In comparing Brazilian and New Zealand university students, Matos and Brown (2015) concluded that Brazilian students had a more negative conception of assessment, speculating that the largely summative use of assessment in higher education was responsible. Using a “draw-a-picture” technique with categorical analysis taken from Brown and Wang (2013), 60% of Brazilian university students' pictures of assessment were dominated by just two themes (i.e. negative emotions and portrayal of assessment as imprecise or inaccurate) (Landim et al. 2015). Hence, independent of data collection method (i.e. document analysis, factor-analysed survey, or free response drawing), the same message appears—assessment is formal and negatively experienced.

Previous research among New Zealand secondary students with Brown's SCoA version 2 inventory reported invariance in responding to the factors according to student sex, year of study, and ethnic group (Hirschfeld and Brown 2009). Reanalysis of the same data found that differences in interest and self-efficacy in reading had no statistically significant effect on how conceptions of assessment related to achievement (Brown and Walton 2017). In contrast, jurisdictional differences in factor means and inter-correlations were observed for the responses of Brazilian university students compared with New Zealand, Hong Kong, and China students on the 6th version of the SCoA inventory (Brown 2013). Furthermore, there were statistically and practically significant differences in SCoA factor means around attendance and effort given on a computer-based low-stakes test at one American university (Wise and Cotten 2009). This suggests demographic and psychological factors within a jurisdiction may properly play little effect on how students respond to the SCoA inventory; however, contextually defined assessment practices may elicit different responses. Hence, an important goal for this Portuguese study was to establish the nature of their responses to the SCoA and discover whether there were demographic factors that influenced responses.

This review reveals a relative lack of insight into university students' perceptions or conceptions of assessment globally let alone in Portugal. As Brown et al. (2009a, p. 9) recognise, “examining students' attitudes towards assessment purposes may help us uncover important factors in what students do before, during, and after assessment”.

The Portuguese context

Portuguese higher education includes both universities and polytechnics. Universities offer solid scientific training, gathering the efforts and competences of teaching and research units, while polytechnics focus on vocational and advanced technical training for the labour market. The Portuguese higher education system includes public higher education, composed of State institutions and foundations, and private higher education, consisting of institutions belonging to private entities and cooperatives (Eurydyce).¹ In Portugal, there are “classical” universities (e.g. University of Lisbon, University of Porto, and University of Coimbra) and the so-called new universities which were founded in the early 1970s, (e.g. University of Minho, University of Aveiro, University of Beira Interior, University of Algarve, University of Trás-os-Montes and Alto Douro, and University of Azores).

Access to higher education in Portugal is gained by national exams at the end of secondary education. In addition to completing a secondary education programme or having a legally equivalent qualification, students must obtain on the entrance examinations established for their intended programme, a classification equal to or higher than the minimum required classification (Decree-law n.º 90/2008). In competitive programmes, for example Medicine and a number of Engineering degrees, the required entry marks are very high (18 or more out of 20). According to the latest official statistics (*Direcção Geral de Estatísticas da Educação e Ciência* 2017) 361,943 students were enrolled in higher education in all study cycles in 2017, with 65% in university and 35% in polytechnics. More than four-fifths (83.6%) of the university students were enrolled in public higher education and 16.4% in private higher education.² In public higher education, students have to pay a tuition fee. The amount of tuition is fixed according to the nature of the programmes and their quality, with a minimum value corresponding to 1.3 of the national minimum salary.³

Portugal has adopted the Bologna Process (Bologna Declaration 1999) which has shaped higher education landscape in the light of its principles and goals within the European Higher Education Area (EHEA). The resulting system provides degrees of comparable value in order to promote mobility, employability, quality assurance, and the development of *life-long learning* (European Commission 2006). Relevant to this study, Decree-law no. 42/2005 stipulates that training be centred on the competencies to be developed by the students, and the centrality of the student in his/her process of training and learning. Students are expected to engage in project-led work, individual study, assessment, and other activities focused on artistic, sociocultural, and sports dimensions. The pace and priorities of the reforms consistent with the Bologna process vary from country to country (Pereira et al. 2016a).

In Portugal, both public and private universities are required to report on their implementation of the Bologna Process. Institutions reported on changes in the teaching and learning process, based on a student-centred learning that promotes autonomous work and student co-responsibility for learning (see reports on the implementation of the Bologna process). Given that assessments are used to select students for entry to university and that universities are expected to implement formative approaches to assessment, or at least to use diverse forms of assessment along with active methodologies in teaching in the light of the Bologna framework,

¹ https://eacea.ec.europa.eu/national-policies/eurydice/content/first-cycle-programmes-60_en

² [http://www.dgeec.mec.pt/np4/estatglobal/%7B\\$clientServletPath%7D/?newsId=308&fileName=DGEEC_DSEE_DEES_2018_EEF_2017_Principais.pdf](http://www.dgeec.mec.pt/np4/estatglobal/%7B$clientServletPath%7D/?newsId=308&fileName=DGEEC_DSEE_DEES_2018_EEF_2017_Principais.pdf)

³ <https://www.dges.gov.pt/pagina/propinas>

we expect students' conceptions of assessment to endorse the formative purposes, but not reject the summative uses of assessment. This appears to suggest that we expect students to illogically endorse inconsistent conceptions; instead, we suggest that student conceptions of assessment will be ecologically rational (Rieskamp and Reimer 2007). In other words, because assessments fulfil these two different functions with important consequences for students (i.e. improved learning outcomes and certification or selection), we expect students to endorse both purposes. Nevertheless, in light of the assessment for learning and Bologna processes, we do anticipate that the formative purposes of assessment will be endorsed more than the summative ones.

Method

The study

This study is part of a larger research project “Assessment in higher education: the potential of alternative methods” funded by the Portuguese Foundation for Science and Technology (Government Funding Agency) (PTDC/MHCCED/2703/2014) and co-financed by European Regional Development Funds (FEDER), Competitiveness and Internationalization Operational Program (POCI-01-0145-FEDER-007562). This broader research project has a strong empirical component in order to identify and analyse assessment practices in five Portuguese public universities from the point of view of both university teachers and students. However, this paper aims (1) to look at students' conceptions of assessment in the same five Portuguese public universities and (2) evaluate student and institutional factors impinging upon how Portuguese students respond to the questionnaire “Students Conceptions of Assessment – version VI” inventory (Brown 2008).

Procedures

In accordance with procedures approved by the Ethics Committee of the University of Minho (Subcommittee on Ethics for Social and Human Sciences, ref. SECSH035 / 2016), programme coordinators and directors were asked for permission to collect data. The survey was disseminated to different programmes in each of the fields of knowledge included in the study (for instance, Medical and Health Sciences included Medicine and Nursing programmes). However, the number of students in those programmes is not publically available. Teachers in the five universities were asked to schedule a time and place in which their students could be invited to complete the questionnaires. Data were collected in person and via email from February to July 2017. In total, 5549 students from five Portuguese public universities participated in the study (Table 1). Out of the 5549 students, 5407 completed the questionnaire in the classroom and 142 students completed the questionnaire replying to the questions by using the link provided via e-mail. Missing data analysis and estimation was carried out to ensure valid inferences would be made. All participants ($n = 24$ deletions, 0.43%) with more than 3 missing values on SCoA items (i.e. $> 10\%$) were deleted. Expectation maximisation was used to impute missing values for the remaining cases. Little's MCAR test chi-square ratio to df was statistically not significant ($\chi^2_{(4144)} = 4980$, $\chi^2/df = 1.2$, $p = 0.27$), suggesting the distribution of missing responses was random and imputation was valid. The final retained sample with no missing values was $N = 5525$.

Table 1 Portuguese higher education participants

Demographic characteristics	<i>N</i>	Percentage (%)
University ^a		
A	2032	36.8
B	620	11.2
C	1351	24.5
D	1110	20.1
E	412	7.5
Gender		
Male	1920	34.8
Female	3488	63.1
Missing	117	2.1
Field of knowledge		
Medical and Health Sciences	1360	24.6
Exact Sciences	210	3.8
Engineering and Technology	1514	27.4
Social Sciences	2000	36.2
Humanities	441	8.0
Cycle of study		
Undergraduate	2807	50.8
Master degree	196	3.5
Integrated master degree	2522	45.6
Year of study		
First year	1774	32.1
Second year	1527	27.6
Third year	1421	25.7
Fourth year	467	8.5
Fifth year	239	4.3
Sixth year	93	1.7
Missing	3	0.1
Total	5525	100

^a The low return rate in universities B and E may be explained, to a great extent, to their size and also to the number of students enrolled in each of the programmes included in this study.

Participants

Approximately two-thirds of the participants were women ($n = 3488$, 63.1%). Just over half of the participants were between 20 and 25 years old (55.7%). Participants were enrolled in different academic programmes in different fields of knowledge: Medical and Health Sciences, Exact Sciences, Engineering and Technology, Social Sciences, and Humanities. About four-fifths of the participants were in the first 3 years of an undergraduate or integrated masters' degree.⁴

Instrument—students conceptions of assessment inventory

Research with the Student Conceptions of Assessment (SCoA) inventory (Brown 2008) into the conceptions of higher education students have about assessment has focused on how students conceive of the purpose, nature, and function of assessment (i.e. it is for improvement,

⁴ An integrated masters' degree (e.g. Engineering and Medicine) differs from the consecutive model (for instance, a 3-year bachelor degree + 2-year masters' degree). The integrated model means that students follow the integrated masters' degree programme from the very beginning.

it predicts the student's future, it improves classroom climate, and it is irrelevant). The SCoA aggregates items into four major inter-correlated factors, with subordinate factors for the improvement and irrelevance factors. The four inter-correlated constructs are (i) assessment improves learning and teaching (improvement); (ii) assessment relates to external factors (external); (iii) assessment has affective benefit (affect); and (iv) assessment is irrelevant (irrelevance).

A Portuguese version of the SCoA-VI inventory was used in Brazil (Matos 2010) and this was the version used in this study. In accordance with Matos (2010), the item (sf4, assessment tells my parents how much I've learnt) was deleted because most university students no longer have their grades reported to their parents. Instead of the positively packed rating scale used in the SCoA-VI, this study adopted Likert's five-point balanced agreement rating scale (i.e. strongly disagree, disagree, neither agree nor disagree, agree, and strongly agree), following from a previous Portuguese study with nursing teachers (Gonçalves 2012). Note, while completely conventional, the restricted number of options in the positive and negative ranges usually reduces variance in the data, potentially making modelling somewhat more complicated.

Studies with New Zealand secondary students and university students (Brown 2013; Brown et al. 2009c) suggest that the hierarchical statistical model fits both samples. However, studies with samples outside New Zealand indicate that no single statistical model exists to capture student beliefs (Brown 2013). Research in Brazil (Brown 2013) has suggested that the hierarchical elements of the SCoA do not work, but Matos and Brown (2015) found that the items aggregated, as per design, into the eight original factors of the SCoA in an inter-correlated model. Bifactor analysis of Brazilian and New Zealand university student SCoA responses (Matos et al. 2019) found full metric equivalence when a general SCoA factor was supplemented with three unique factors (i.e. assessment improves classroom climate, assessment improves teaching and learning, and assessment is irrelevant). These Brazil-New Zealand comparative studies suggest that there may be universal aspects of how students conceive of assessment and support the use of the inventory in Portugal.

This study surveys Portuguese university students to test the validity of the SCoA for use in a second Portuguese language context and to test for differences between groups of students based on membership in relevant educational categories (i.e. their university, year of study, field of knowledge, and cycle of study). Thus, this study contributes to a growing understanding of global and local aspects of student experience of higher education assessment.

Analysis

Because multiple previously published statistical models for the SCoA inventory existed, four models were tested with confirmatory factor analysis to identify the model that best corresponded with the data. The models were:

1. Brown's NZ SCoA-VI model (Brown et al. 2009c),
2. Matos' eight-factor inter-correlational model (Matos and Brown 2015),
3. A bifactor model with a general factor and four main factors (Weekers et al. 2009), and
4. An inter-correlated four main factors model, in which all subordinate factors are removed so that items load directly on the higher order factor.

Model fit was examined in line with conventional practices (Fan and Sivo 2007; Marsh et al. 2004). A model need not be rejected if:

- χ^2 per degree of freedom was statistically non-significant,
- Gamma hat and comparative fit index (CFI) > .90,
- Root mean square error of approximation (RMSEA) < .08, with 90% confidence interval being less than 0.08, and
- Standardized root mean residual (SRMR) < .08.

When choosing between competing models, differences in the Akaike information criterion (Δ AIC) greater than 10 indicate that the smaller value is closer to the data (Burnham and Anderson 2004). By taking advantage of features of the data, it is possible to improve the fit of a model by inspection of the modification indices. Items with strong attraction to other factors violate the assumption of simple structure in which items clearly belong only to one factor. Deletion of those items can improve the fit but such actions have to be theoretically justifiable.

Once a model is selected, the stability of the model across participant demographic groupings was conducted using nested multigroup confirmatory factor analysis (Brown et al. 2017). Nested invariance testing examines the equivalence of regression weights (metric equivalence) and item intercepts (i.e. scalar invariance). When differences in the CFI are < .01, the addition of equivalence constraints is considered not to have changed the quality of fit. This approach first fixes regression weights to be equivalent, before testing equivalence of intercepts or starting points. All confirmatory factor analysis and invariance testing were carried out in Amos v25 (IBM 2017) using maximum likelihood estimation.

Comparison of mean scores usually requires that the groups have statistically equivalent item regressions and intercepts. To evaluate the differences in demographic categories, the Bartlett factor scores were calculated for each scale or factor in the model. Bartlett factor scores are latent scores derived from the weighted contribution of items within each factor (DiStefano et al. 2009). The scores are centred on zero with a mean of 1, meaning they behave like a *z*-score.

Results

Model selection

The fit indices for the four tested models are shown in Table 2. According to AIC value, model 2 with eight inter-correlated factors was best fitting (Δ AIC = 592 over model 3). Because the chi-squared per degree of freedom of model 2 was > 17, the support for this model was weak. However, because the chi-square test is very sensitive to big sample size (Wheaton et al. 1977), a randomly selected sample of 500 participants (model 2A) was drawn and tested against the same model. This produced a statistically not significant result for the chi-square per degree of freedom ratio indicating that this model did not need to be rejected.

While 13 items had substantial sum of modification indices, just four items (i.e. two in classroom environment [ce6, ce5] and two in irrelevance [ig3, bd5]) had to be removed to achieve excellent model fit (model 6 in Table 2). Figure 1 shows the structure of the trimmed model 2 showing eight inter-correlated factors, with 28 items. Note that three factors only have two items which are normally unacceptable; however, in a multi-factorial inventory, factors can be estimated with just 2 items (Bollen 1989).

Multiple-group confirmatory factor analysis

A series of multigroup confirmatory factor analyses (MG-CFA) were conducted with grouping variables being (a) field of knowledge, (b) year, (c) university, and (d) cycle of study. Model fit indices (Table 3) show that all four approaches yielded similar fit. However, the model based on cycle of study had the smallest AIC value ($\Delta AIC > 200$) suggesting that it should be used to analyse factor scores in trimmed model 2.

Before comparing the mean scores, invariance testing of cycle of study's three groups was conducted within MG-CFA. Each cycle group was found to separately fit the data well and the fit improved when the three groups were combined in one model (Table 4). The metric and scalar equivalence tests were sequentially $\Delta CFI < 0.01$ indicating equivalence in regression weights and intercepts between the three cycle groups. Thus, factor loadings and means can be compared between these three groups.

Factor mean scores and correlations

Mean scores for the factors, taking the average of all items belonging to the factor, were highest for Student Improvement, close to neutral for five scales, and lowest for Ignore (Table 5). Differences in mean scores were relatively trivial to small, except Student Improvement which had large positive effects relative to all other factors (mean Cohen's $d = 1.83$) and Ignore which had large negative effects relative to all other factors (mean Cohen's $d = 1.87$). The small variations in mean scores among six factors suggest that, after endorsing assessment for improving student learning and rejecting the act of ignoring assessment, there was a relatively neutral stance towards assessments for accountability, improvement, and socio-affective purposes. Thus, the students' conceptions can most succinctly be summarised as assessments helped them improve and were not ignored.

Table 5 also shows the factor inter-correlation matrix for trimmed model 2. It shows that there are potentially two major dimensions within the eight factors. As previously reported, the two accountability factors (i.e. Student Future and School Quality), the two improvement factors (i.e. Student and Teacher), and the two emotional factors (i.e. Personal Enjoyment and Class Environment) are all positively inter-correlated with each other. In contrast, and consistent with previous studies, the two irrelevance factors (i.e. Bad and Ignore) are negatively

Table 2 SCoA measurement model fit indices for four comparative models

Model	# factors	χ^2 (df)	χ^2/df	CFI	Gamma hat	RMSEA	90%CI	SRMR	AIC
1. Brown	10	9107 (451)	20.16***	.817	.91	.059	.058–.060	.060	9325
2. Matos and Brown	8	7799 (436)	17.89***	.844	.92	.055	.054–.056	.053	7983
3. Model 2 random $n = 500$	8	1228 (436)	2.82 ^{ns}	.806	.91	.060	.056–.064	.067	1412
4. Bifactor	9	8383 (432)	19.41***	.832	.92	.058	.057–.059	.052	8575
5. Four main	4	13,976 (458)	30.51***	.714	.87	.073	.072–.074	.079	14,180
6. Model 2 trimmed	8	4824 (322)	14.98***	.88	.95	.050	.049–.052	.043	4992

CFI, comparative fit index; RMSEA, root mean square error of approximation; SRMR, standardised root mean residual; AIC, Akaike information criterion; *** $p < .001$; ns, not significant

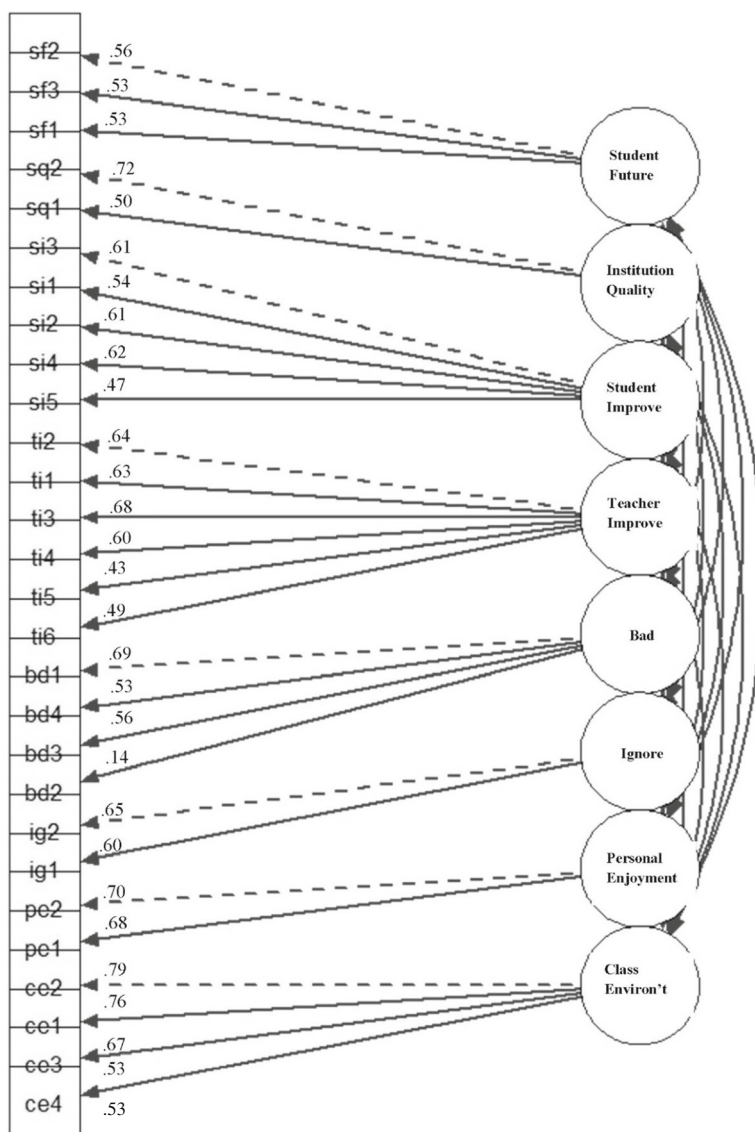


Fig. 1 Trimmed model 2 based on Matos and Brown (2015) for Portugal university students. Note: All values are standardised; diagram from Jamovi 0.9 using Lavaan package in R

Table 3 Model fit indices by different grouping methods on trimmed Model 2

Model	χ^2 (df)	χ^2/df	CFI	Gamma hat	RMSEA	90%CI	SRMR	AIC
University	6406 (1610)	3.979*	.869	.94	.023	.023–.024	.048	7526
Year	6815 (1932)	3.527 ^{ns}	.864	.94	.021	.021–.022	.045	8159
Field of knowledge	6458 (1610)	4.011*	.867	.94	.023	.023–.024	.050	7298
Cycle of study	6627 (1078)	6.148*	.848	.93	.031	.030–.031	.048	7075

CFI, comparative fit index; RMSEA, root mean square error of approximation; SRMR, standardised root mean residual; AIC, Akaike information criterion; * $p < .05$; ^{ns}, not significant

Table 4 MG-CFA fit indices for separate and nested invariance testing based on cycle of study

Group	Fit statistics						
	χ^2 (df)	χ^2/df	CFI	Gamma hat	RMSEA	90%CI	SRMR
Separate							
Undergraduate	2706 (322)	8.407**	.876	.94	.051	.050–.053	.046
Integrated master	2434 (322)	7.560**	.867	.94	.051	.049–.053	.045
Master	541 (322)	1.681 ^{ns}	.831	.93	.059	.050–.068	.064
Combined							
Unconstrained	6627 (1078)	6.148*	.848	.93	.031	.030–.031	.048
Measurement weights	6647 (1098)	6.054*	.848	.93	.030	.030–.031	.047
Measurement intercepts	6764 (1126)	6.008*	.845	.94	.029	.029–.031	.047

CFI, comparative fit index; RMSEA, root mean square error of approximation; SRMR, standardised root mean residual; AIC, Akaike information criterion; * $p < .05$; ** $p < .01$; ns, not significant

correlated with the six other factors but positively correlated with each other. Interestingly, Bad had a stronger negative correlation with Teacher Improvement, while Ignore had a stronger negative correlation with Student Improvement, suggesting that assessment that informs teachers' teaching is not bad, while assessment that guides student improvement is not ignored.

MANOVA was conducted with factor mean scores as dependent variables, and demographic variables as independent variables (i.e. field of knowledge, year, cycle of study, university, sex) as both main effects and two-way interaction effects. Given the large number of comparisons made, alpha was set at $p \leq .01$. Note also that not all universities provided students in all fields of knowledge (e.g. university 3 had no students in exact sciences, university 5 had no students in social sciences or humanities); this meant the two-way interaction for university by field of knowledge was excluded from the MANOVA. Furthermore, only two of the five fields of knowledge had all three cycles of study, so this interaction was also removed from the MANOVA. Likewise, cycle of study 2 had < 10 students in year 3 onward and there were no cycle of study 1 or 2 students in year 6; hence, the cycle of study by year interaction was also excluded.

Four of the main effects met the threshold for consideration (i.e. field of knowledge, $F_{(32)} = 3.15$, $p < .001$; sex, $F_{(8)} = 6.19$, $p < .001$; year, $F_{(40)} = 2.30$, $p < .001$; and cycle of study, $F_{(16)} = 3.60$, $p < .001$), while five two-way interactions also met that standard (i.e. field of knowledge by year, $F_{(128)} = 1.45$, $p = .001$; and sex, $F_{(32)} = 1.81$, $p = .003$; university by sex, $F_{(32)} = 1.85$,

Table 5 Trimmed model 2 factor descriptive statistics

	Factor mean scores		Inter-correlations						
Factor	<i>M</i>	<i>SD</i>	II.	III.	IV.	V.	VI.	VII.	VIII.
I. Student future	2.81	0.72	.64	.44	.71	– .31	– .18	.67	.42
II. School quality	3.01	0.82		.31	.60	– .24	– .06	.46	.32
III. Student improvement	3.89	0.49			.55	– .21	– .67	.44	.26
IV. Teacher improvement	3.11	0.63				– .55	– .22	.76	.49
V. Bad	3.01	0.59					.29	– .45	– .26
VI. Ignore	1.86	0.66						– .17	– .06
VII. Personal enjoyment	2.76	0.83							.44
VIII. Class environment	3.01	0.76							

$p = .002$; year, $F_{(136)} = 1.42$, $p = .001$; and cycle of study, $F_{(64)} = 2.07$, $p < .001$). Inspection of between-subject univariate tests just for these three main effects and six interactions identified that all factors had at least one predictor with alpha at $p < .01$ (Table 6). However, the total variance explained for each factor fell in the small range ($f^2 < .11$, Cohen 1992), suggesting that these background variables do little to explain how students conceive of assessment. Instead, intra-individual psychological or social factors are more likely to impact these attitude or belief scores. In light of this small power, no further analysis was carried out to explore how student conceptions of assessment varied by demographic conditions.

Discussion

This survey of Portuguese university students with a translated and adapted version of the Student Conceptions of Assessment (version VI) inventory successfully replicated the inter-correlated eight-factor model reported in Brazil (Matos and Brown 2015) after trimming four items in two factors. The eight conceptions related to the purpose of assessment are as follows: (1) predicts student future; (2) evaluates school quality; (3) supports student learning improvement; (4) supports improvement of teaching; (5) is bad; (6) is ignored; (7) is personally enjoyable; and (8) supports positive class environment. This study successfully replicates the previous Brazilian survey of university students (Matos and Brown 2015) and confirms that the items in the SCoA-VI, developed in New Zealand with secondary school students (Brown 2008), have generalizability as measures of Portuguese-speaking university student conceptions of assessment.

The trimmed model was strongly invariant by cycle of study supporting the comparison of factor means across groups. Factor mean score differences existed but were trivially small based on institutional and student demographic characteristics. This contradicted the expectation that differences in experiences related to discipline, degree, university, length of study, or sex would lead to different conceptions of assessment. Overall, students tended to agree with the student improvement factor and reject the ignore concept; otherwise, their conceptions

Table 6 Statistically significant demographic predictors of student conceptions of assessment

SCoA factor	Demographic predictors	Cohen's f^2 effect size
Classroom environment	Field of knowledge; cycle of study; year; university by sex; university by cycle of study; sex by field of knowledge; year by field of knowledge	.10
Personal enjoyment	Field of knowledge; university by cycle of study field of knowledge by cycle of study;	.04
Ignore	Sex; university by sex; university by cycle of study; sex by field of knowledge; year by field of knowledge	.04
Bad	Field of knowledge; year; university by sex	.04
Teacher improvement	Year; field of knowledge by cycle of study; university by sex; university by year	.04
Student improvement	Sex; year; university by cycle of study; year by field of knowledge	.04
School quality	Field of knowledge	.02
Student future	Year	.03

MANOVA design: intercept + field of knowledge + cycle of study + year + sex + university + university $\times$ sex + university $\times$ year + university $\times$ cycle of study + sex $\times$ year + sex $\times$ field of knowledge + sex $\times$ cycle of study + field of knowledge $\times$ year

towards the other factors were relatively neutral. This result echoes previous studies which have suggested the stability of the SCoA factors across both demographic and psychological factors (Brown and Walton 2017; Hirschfeld and Brown 2009). Such stability within a population is a desirable feature of any research instrument. Nonetheless, there may be more substantial differences in SCoA scores if observational data around assessment practices (e.g. effort given, time on task, perseverance in light of failure) had been available. Given the nature of the study, finding common practices across so many different disciplines and universities was beyond the scope of this study.

Elsewhere, it has been shown that academically successful students actively see assessment as a tool for improving their learning and performance (Brown et al. 2009c, 2016) and accept that assessment judges their learning (Brown and Hirschfeld 2008). The positive endorsement of student improvement in this study is entirely consistent with notions that student responses, attitudes, and conceptions of assessment that support self-regulation of learning are appropriate and adaptive for successful learning outcomes (Brown 2011). The relatively negative to weak endorsement for the two affective purposes is consistent with research in New Zealand, which has shown that endorsement of classroom environment purpose had no statistically significant relationship to achievement (Brown et al. 2009c). Furthermore, reliance on peers for feedback has not been shown to contribute to greater performance among New Zealand university students (Brown et al. 2016). Combined with the rejection of the concept that assessment should be ignored, the results suggest that these students have a view that assessment is primarily about paying attention to what they got wrong so that they can improve. These results indicate Portuguese university students are already trying to use whatever information they get to improve their performance, which is what successful, self-regulating students do.

The slightly negative attitude towards the teacher improvement purpose was somewhat surprising relative to results in other jurisdictions (Brown 2011). Higher performance was seen among secondary school students who had confidence that their instructors used assessments to improve instruction (Brown et al. 2009c). However, it may be difficult in higher education for students to perceive that teachers are using assessment for improved teaching. Nonetheless, it would be tragic if the belief in assessment for improvement were left solely to the responsibility of the learner.

It behoves instructors, departments, and institutions to ensure that assessment within courses and programmes is designed to permit the possibility of gaining feedback that supports learning in subsequent assessments. This probably means not relying solely on percentage scores or letter grades to communicate information about what students need to do to improve their work (Brown and Hattie 2012); diagnostic insights are needed which probably come in the form of commentary rather than scores (Lipnevich and Smith 2009). Further, it means that assessments have to scaffold in such a way that feedback from early assessments helps students learn what they need to do on later and summative or terminal tasks. Unfortunately, this is an approach that is not universally implemented in higher education (Winstone and Boud 2019).

The negative and relatively strong correlation between Teacher Improvement and Bad suggests that students accept, like previous studies, that assessments are not bad if they help teachers do a better job. Nonetheless, it may be difficult for higher education students to perceive that instructors actually use assessments to modify or improve their instruction. Given semester-long courses in which all assessment and course details have to be published in advance, it is very difficult for students to notice if teachers make changes based on assessment results. The only tangible evidence of teacher use is likely to appear in the kind, speed, and

quality of feedback provided. Furthermore, the use of national summative exams to provide students access to degrees and institutions is likely to lead students to a somewhat narrow view of assessment; how can teachers be using it for improvement when it is about my terminal performance with little or no feedback beyond a score?

It is an unfortunate but understandable limitation of this study that a measure of academic performance could not be included, making this a matter for future investigations to test the claim that students who endorse student improvement and reject ignoring it have higher grades. Furthermore, future research should incorporate the SCoA with measures of self-regulation to ascertain how self-reported self-regulation relates to these eight conceptions of assessment. Embedding such research in or around actual assessment tasks would also be highly desirable (e.g. Brown et al. 2016; Peterson et al. 2015) to identify characteristics of assessments that support self-regulating conceptions of assessment.

In conclusion, it would appear the Portuguese university students have an overall conception of assessment that expresses personal responsibility to use insights from assessment to improve their learning, performance, and achievement. The success of the Bologna Process in terms of how assessment is experienced by students is not strongly evident in these five institutions. Nonetheless, the ambition of the Bologna Process is a positive one and further efforts are needed by universities to mitigate an approach to assessment that is purely summative.

Acknowledgements Statistical assistance was provided by Anran Zhao, a doctoral candidate at The University of Auckland, with funding from the Faculty of Education and Social Work.

Funding information The project reported in this paper was funded by the Portuguese Foundation for Science and Technology (Government Funding Agency) (PTDC/MHCCED/2703/2014) and co-financed by European Regional Development Funds (FEDER), Competitiveness and Internationalization Operational Program (POCI-01-0145-FEDER-007562).

References

- Barreira, C., Bidarra, G., Monteiro, F., Vaz-Rebelo, P., & Alfêres, V. (2017). Avaliação das aprendizagens no ensino superior. Perceções de professores e estudantes nas universidades portuguesas [Assessment in higher education: perceptions of university teachers and students in Portuguese universities]. *Revista Iberoamericana de Educación Superior*, 8(21), 24–36.
- Biggs, J. (2003). *Teaching for quality learning at university*. Buckingham: SHRE and Open University Press.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Boud, D. (1995). Assessment and learning: contradictory or complementary? In P. Knight (Ed.), *Assessment for learning in higher education* (pp. 35–48). London: Kogan Page.
- Boud, D., & Falchikov, N. (2007). *Rethinking assessment in higher education: learning for the long term*. New York: Routledge.
- Brown, G. T. L. (2008). *Students' conceptions of assessment inventory (SCoA Version VI) [Measurement instrument]*. Auckland: University of Auckland. <https://doi.org/10.17608/k6.auckland.4596820.v1>.
- Brown, G. T. L. (2011). Self-regulation of assessment beliefs and attitudes: a review of the students' conceptions of assessment inventory. *Educational Psychologist*, 31(6), 731–748.
- Brown, G. T. L. (2013). Student conceptions of assessment across cultural and contextual differences: university student perspectives of assessment from Brazil, China, Hong Kong, and New Zealand. In G. A. D. Liem & A. B. I. Bernardo (Eds.), *Advancing cross-cultural perspectives on educational psychology: a festschrift for Dennis McInerney* (pp. 143–167). Charlotte: Information Age Publishing.
- Brown, G. T. L., & Harris, L. R. (2012). Student conceptions of assessment by level of schooling: further evidence for ecological rationality in belief systems. *Australian Journal of Educational and Developmental Psychology*, 12, 46–59.

- Brown, G. T. L., & Hattie, J. A. (2012). The benefits of regular standardized assessment in childhood education: guiding improved instruction and learning. In S. Suggate & E. Reese (Eds.), *Contemporary debates in childhood education and development* (pp. 287–292). London: Routledge.
- Brown, G. T. L., & Hirschfeld, G. (2008). Students' conceptions of assessment: links to outcomes. *Assessment in Education: Principles, Policy & Practice*, 15(1), 3–17.
- Brown, S., & Knight, P. (1994). *Assessing learners in higher education*. London: Kogan Page.
- Brown, G. T. L., & Walton, K. F. (2017). The effect of conceptions of assessment upon reading achievement: an evaluation of the influence of self-efficacy and interest. *Interdisciplinary Education and Psychology*, 1(103). <https://doi.org/10.31532/InterdiscipEducPsychol.1.1.003>.
- Brown, G. T. L., & Wang, Z. (2013). Illustrating assessment: How Hong Kong university students conceive of the purposes of assessment. *Studies in Higher Education*, 38(7), 1037–1057. <https://doi.org/10.1080/03075079.2011.616955>.
- Brown, G. T. L., & Wang, Z. (2016). Understanding university student conceptions of assessment: cultural similarities and jurisdictional differences between Hong Kong and China. *Social Psychology of Education*, 19, 151–173.
- Brown, G., Bull, J., & Pendlebury, M. (1997). *Assessing student learning in higher education*. London: Routledge.
- Brown, G. T. L., Irving, S. E., & Peterson, E. R. (2009a). *The more I enjoy it the less I achieve: the negative impact of socio-emotional purposes of assessment and feedback on academic performance*. Paper presented at EARLI 2009 biennial conference, Amsterdam, NL.
- Brown, G. T. L., Irving, S. E., Peterson, E. R., & Hirschfeld, G. H. F. (2009b). Use of interactive-informal assessment practices: New Zealand secondary students' conceptions of assessment. *Learning and Instruction*, 19(2), 97–111. <https://doi.org/10.1016/j.learninstruc.2008.02.003>.
- Brown, G. T. L., Peterson, E., & Irving, S. (2009c). Self-regulatory beliefs about assessment predict mathematics achievement. In D. M. McInerney, G. T. L. Brown, & G. A. D. Liem (Eds.), *Students perspectives on assessment: what students can tell us about assessment for learning* (pp. 159–186). Charlotte: Information Age.
- Brown, G. T. L., Pishghadam, R., & Shayesteh Sadafian, S. (2014). Iranian university students' conceptions of assessment. *Assessment Matters*, 6, 5–33.
- Brown, G. T. L., Peterson, E. R., & Yao, E. S. (2016). Student conceptions of feedback: impact on self-regulation, self-efficacy, and academic achievement. *British Journal of Educational Psychology*, 86(4), 606–629. <https://doi.org/10.1111/bjep.12126>.
- Brown, G. T. L., Harris, L. R., O'Quin, C., & Lane, K. E. (2017). Using multi-group confirmatory factor analysis to evaluate cross-cultural research: identifying and understanding non-invariance. *International Journal of Research & Method in Education*, 40(1), 66–90. <https://doi.org/10.1080/1743727X.2015.1070823>.
- Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods & Research*, 33(2), 261–304. <https://doi.org/10.1177/0049124104268644>.
- Chen, J., & Brown, G. T. L. (2018). Chinese secondary school students' conceptions of assessment and achievement emotions: endorsed purposes lead to positive and negative feelings. *Asia Pacific Journal of Education*, 38(1), 91–109. <https://doi.org/10.1080/02188791.2018.1423951>.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159.
- De Oliveira, J., & Flores, M. A. (2017). Conceções e práticas avaliativas no Ensino Superior: aprendendo com os alunos [Conceptions and practices of assessment in Higher Education: learning with students]. In M. A. Flores, M. P. Alves, E. A. Machado, & S. Fernandes (Orgs.), *Atas do I Congresso Internacional sobre Avaliação das Aprendizagens e Sucesso Escolar* (pp. 9–14). Braga: Centro de Investigação em Estudos da Criança (CIEC). ISBN 978-972-8952-45-7.
- Decree-Law n° 42/2005. (2005). *Ministry of Science, Innovation and higher education*.
- DiStefano, C., Zhu, M., & Mindrila, D. (2009). Understanding and using factor scores: considerations for the applied researcher. *Practical Assessment, Research & Evaluation*, 14(20) <http://pareonline.net/getvn.asp?v=14&n=20>.
- Dochy, F., & McDowell, L. (1997). Assessment as a tool for learning. *Studies in Educational Evaluation*, 23(4), 279–298.
- European Commission, Directorate General Education and Culture. (2006). *The extent and impact of higher education curricular reform across Europe: final report to the Directorate-General for Education and Culture of the European Commission*.
- Fan, X., & Sivo, S. A. (2007). Sensitivity of fit indices to model misspecification and model types. *Multivariate Behavioral Research*, 42(3), 509–529. <https://doi.org/10.1080/00273170701382864>.
- Fernandes, D. (2009). Educational assessment in Portugal. *Assessment in education: Principles, Policy & Practice*, 16(2), 227–247.

- Fernandes, D. (2014). Práticas de ensino e de avaliação de docentes de quatro universidades portuguesas [Practices of teaching and assessment in four Portuguese universities]. In D. Fernandes, A. Borralho, C. Barreira, A. Monteiro, D. Catani, E. Cunha, & P. Alves (Eds.), *Avaliação, ensino e aprendizagens em Portugal e no Brasil: realidades e perspectivas* (pp. 97–135). Lisboa: Educa.
- Fernandes, D. (2015). Pesquisa de percepções e práticas de avaliação no ensino universitário português [Review of perceptions and practices of assessment in higher education]. *Estudos em Avaliação Educacional*, 26(63), 596–629.
- Fletcher, R., Meyer, L., Anderson, H., Johnston, P., & Rees, M. (2012). Faculty and students conceptions of assessment in higher education. *Higher Education*, 64(1), 119–133.
- Flores, M., Veiga Simão, M., Barros, A., & Pereira, D. (2015). Perceptions of effectiveness, fairness and feedback of assessment methods: a study in higher education. *Studies in Higher Education*, 40(9), 1523–1534.
- Gonçalves, R. (2012). *Avaliação das Aprendizagens dos Estudantes de Enfermagem em Ensino Clínico: das concepções e práticas docentes à apresentação de propostas de intervenção* [Assessing student learning in nursing in clinical teaching: from conceptions and practices to proposals of intervention] (Unpublished doctoral thesis). University of Coimbra, Coimbra, Portugal.
- Gulikers, J., Kester, L., Kirschner, P., & Bastiaens, T. (2008). The effect of practical experience on perceptions of assessment authenticity, study approach, and learning outcomes. *Learning and Instruction*, 18, 172–186.
- Hirschfeld, G. H. F., & Brown, G. T. L. (2009). Students' conceptions of assessment: factorial and structural invariance of the SCoA across sex, age, and ethnicity. *European Journal of Psychological Assessment*, 25(1), 30–38. <https://doi.org/10.1027/1015-5759.25.1.30>.
- IBM. (2017). *AMOS (Version 25.0.0)* [computer program]. Wexford: Amos Development Corporation.
- Landim, M. V., Matos, D. A. S., & Brown, G. T. L. (2015). A avaliação sob o ponto de vista dos estudantes: O uso de desenhos para a análise de concepções de avaliação [Assessment from the point of view of students. The use of drawing to analyse conceptions of assessment]. *Psicologia, Educação e Cultura*, 19(2), 55–74.
- Lipnevich, A. A., & Smith, J. K. (2009). The effects of feedback on student examination performance. *Journal of Experimental Psychology: Applied*, 15, 319–333.
- 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. https://doi.org/10.1207/s15328007sem1103_2.
- Matos, D. (2010). *A avaliação no ensino superior: concepções múltiplas de estudantes Brasileiros* [Assessment in higher education: multiple conceptions of Brazilian students] (Unpublished doctoral thesis). Universidade Federal de Minas Gerais, Brasil.
- Matos, D. A. S., & Brown, G. T. L. (2015). Comparing university student conceptions of assessment: Brazilian and New Zealand beliefs. In C. Carvalho & J. Conboy (Eds.), *Feedback, identidade, trajetórias escolares: Dinâmicas e consequências* (pp. 177–194). Lisbon: Universidade de Lisboa, Instituto de Educação.
- Matos, D., Cirino, S., & Brown, G. (2009). Students' conceptions of assessment in higher education in Brazil. In D. M. McInemey, G. T. L. Brown, & G. A. D. Liem (Eds.), *Student perspectives of assessment. What students can tell us about assessment for learning* (pp. 235–253). North Carolina: Information age.
- Matos, D. A. S., Brown, G. T. L., & Cirino, S. D. (2012). Concepções de avaliação de alunos universitários: uma revisão da literatura [Conceptions of assessment of students in higher education: a literature review]. *Estudos em Avaliação Educacional*, São Paulo, 23(52), 204–231.
- Matos, D. A. S., Cirino, S. D., Brown, G. T. L., & Leite, W. L. (2013). A avaliação no ensino superior: Concepções múltiplas de estudantes Brasileiros [Assessment in higher education: Multiple conceptions of Brazilian students]. *Estudos em Avaliação Educacional*, 24(54), 172–193. <https://doi.org/10.18222/eaec245420131907>.
- Matos, D. A. S., Brown, G. T. L., & Gomes, C. (2019). Bifactor invariance analysis of student conceptions of assessment inventory. *Psico-USF*, 24(4).
- McMillan, J. H. (2016). Section discussion: student perceptions of assessment. In G. T. L. Brown & L. R. Harris (Eds.), *Handbook of human and social conditions in assessment* (pp. 221–243). New York: Routledge.
- Myers, C., & Myers, S. (2015). The use of learner-centered assessment practices in the United States: the influence of individual and institutional contexts. *Studies in Higher Education*, 40(10), 1904–1918.
- Novak, E., & Johnson, T. (2012). Assessment of students' emotions in game-based learning. In D. Ifenthaler, D. Eseryel, & X. Ge (Eds.), *Assessment in game-based learning: foundations, innovations, and perspectives* (pp. 379–399). New York: Springer.
- Otunuku, M., Brown, G. T. L., & Airini. (2013). Tongan secondary students' conceptions of schooling in New Zealand relative to their academic achievement. *Asia Pacific Education Review*, 14(3), 345–357. <https://doi.org/10.1007/s12564-013-9264-y>.

- Panadero, E., Fraile, J., Fernández Ruiz, J., Castilla-Estévez, D., & Ruiz, M. A. (2018). Spanish university assessment practices: examination tradition with diversity by faculty. *Assessment & Evaluation in Higher Education*. Online first. <https://doi.org/10.1080/02602938.2018.1512553>.
- Pereira, D. (2016). *Assessment in higher education and quality of learning: perceptions, practices and implications* (Unpublished doctoral thesis). University of Minho, Braga, Portugal.
- Pereira, D., Flores, M., & Niklasson, L. (2016a). Assessment revisited: a review of research in assessment and evaluation in higher education. *Assessment & Evaluation in Higher Education*, 41(7), 1008–1032.
- Pereira, D., Flores, M., Veiga Simão, A., & Barros, A. (2016b). Effectiveness and relevance of feedback in higher education: a study of undergraduate students. *Studies in Educational Evaluation*, 49, 7–14.
- Pereira, D., Flores, M., & Barros, A. (2017a). Perceptions of Portuguese undergraduate students about assessment: a study in five public universities. *Educational Studies*, 43, 442–463.
- Pereira, D., Niklasson, L., & Flores, M. A. (2017b). Students' perceptions of assessment: a comparative analysis between Portugal and Sweden. *Higher Education*, 73(1), 153–173.
- Peterson, E., & Irving, E. (2008). Secondary school students' conceptions of assessment and feedback. *Learning and Instruction*, 18, 238–250.
- Peterson, E. R., Brown, G. T. L., & Jun, M. C. (2015). Achievement emotions in higher education: a diary study exploring emotions across an assessment event. *Contemporary Educational Psychology*, 42(1), 82–96. <https://doi.org/10.1016/j.cedpsych.2015.05.002>.
- Pitt, E., & Norton, L. (2017). Now that's the feedback I want!' Students' reactions to feedback on graded work and what they do with it. *Assessment & Evaluation in Higher Education*, 41, 1–18.
- Race, P. (1995). What has assessment done for us - and to us? In P. Knight (Ed.), *Assessment for learning in higher education* (pp. 61–74). London: Kogan Page.
- Rieskamp, J., & Reimer, T. (2017). Ecological rationality. In R. F. Baumeister & K. D. Vohs (Eds.), *Encyclopedia of social psychology* (pp. 273–275). Thousand Oaks: Sage.
- Ryan, R., & Henderson, M. (2018). Feeling feedback: students' emotional responses to educator feedback. *Assessment & Evaluation in Higher Education*, 43(6), 880–892.
- Segers, M., & Tillema, H. (2011). How Dutch secondary teachers and students conceive the purpose of assessment? *Studies in Educational Evaluation*, 37, 49–54.
- Solomonidou, G., & Michaelides, M. (2017). Students' conceptions of assessment purposes in a low stakes secondary-school context: a mixed methodology research. *Studies in Educational Evaluation*, 52, 35–41.
- Struyven, K., & Devesa, J. (2016). Students' perceptions of novel forms of assessment. In G. T. L. Brown & L. R. Harris (Eds.), *Handbook of human and social conditions in assessment* (pp. 129–144). New York: Routledge.
- Struyven, K., Dochy, F., & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: a review. *Assessment & Evaluation in Higher Education*, 30(4), 331–347.
- Wang, Z., & Brown, G. (2014). Hong Kong tertiary students' conceptions of assessment of academic ability. *Higher Education Research and Development*, 33(5), 1063–1077. <https://doi.org/10.1080/07294360.2014.890565>.
- Webber, K. (2012). The use of learner-centered assessment in US colleges and universities. *Research in Higher Education*, 53(2), 201–228.
- Weekers, A. M., Brown, G., & Veldkamp, B. P. (2009). Analyzing the dimensionality of the students' conceptions of Assessment inventory. In D. M. McInerney, G. T. L. Brown, & G. A. D. Liem (Eds.), *Student perspectives on assessment: What students can tell us about assessment for learning* (pp. 135–157). Charlotte: Information Age Publishing.
- Wheaton, B., Muthén, B., Alwin, D. F., & Summers, G. F. (1977). Assessing reliability and stability in panel models. *Sociological Methodology*, 8, 84–136. <https://doi.org/10.2307/270754>.
- Winstone, N., & Boud, D. (2019). Exploring cultures of feedback practice: the adoption of learning-focused feedback practices in the UK and Australia. *Higher Education Research and Development*, 38(2), 411–425. <https://doi.org/10.1080/07294360.2018.1532985>.
- Wise, S. L., & Cotten, M. R. (2009). Test-taking effort and score validity: the influence of student conceptions of assessment. In D. M. McInerney, G. T. L. Brown, & G. A. D. Liem (Eds.), *Student perspectives on assessment: what students can tell us about assessment for learning* (pp. 187–205). Charlotte: Information Age Publishing.
- Zimmerman, B. (2008). Investigating self-regulation and motivation: historical background, methodological developments, and future prospects. *American Educational Research Journal*, 45(1), 166–183.