ISSN 2190-0493 (print) | ISSN 2190-0507 (digital)

Psychological Test and Assessment Modeling

Volume 63, 2-2021

Special issue: Assessment of Educational and Learning Capital

Guest-Editors: Alejandro Veas & Albert Ziegler

Pabst Science Publishers | Lengerich/Westfalen, Germany

Psychological Test and Assessment Modeling

Volume 63, 2021 (2)

Special issue: Assessment of Educational and Learning Capital
Guest-Editors: Alejandro Veas & Albert Ziegler

CONTENTS

Psychological Test and Assessment Modeling

formerly Psychologische Beiträge/Psychology Science Quarterly

Editor:
Klaus D. Kubinger
c/o Faculty of Psychology
University of Vienna
Liebiggasse 5
A-1010 Vienna (Austria)
klaus.kubinger@univie.ac.at

Scientific Office Manager:
Bettina Hagenmüller
p-t-a-m@univie.ac.at

Technical Support:
Jan Steinfeld
jan.steinfeld@univie.ac.at

PABST SCIENCE PUBLISHERS
Eichengrund 28
D-49525 Lengerich
Tel. +49 (0) 5484 – 308
Fax +49 (0) 5484 – 550
Postgiro Kto. Köln 208157-503
www.pabst-publishers.com
www.psychologie-aktuell.com

Publishers Office: Silke Haarlammert
haarlammert@pabst-publishers.com

Producer: Patrick Orths
orths@pabst-publishers.com

Subscription price per year:
(4 issues) 60,- EUR/USD, incl. postage and VAT, single copy 15,- EUR/USD, incl. postage and VAT.

ISSN 2190-0493 (Print)
ISSN 2190-0507 (Internet)
www.psychological-test-and-assessment-modeling.com

Psychometric perspectives in Educational and Learning Capitals: Development and validation of a scale on student evaluation of teaching in Higher Education

Tarquino Sánchez[a], Alejandro Veas[b], Raquel Gilar-Corbí[b] & Juan-Luis Castejón[b]

[a] National Polytechnic School of the Ecuador. Department of Electronic, Telecommunications and information networks

[b] University of Alicante. Department of Developmental Psychology and Didactics

Abstract:

Student evaluation of teaching is an important topic in the field of education, and different rating scales have been developed in international contexts. However, there have been methodological problems that may lead to ineffective construct measurement. The aim of this study Considering as an extension of the Actiotope Model of Giftedness, the aim of this study was to support an internal structure of a new 32 item instrument in a large public university in Ecuador, using a sample of 6110 students. Data were analysed based on the item response theory, including unidimensional and multidimensional Rasch rating scale models to examine three theory-based constructs. Preference was given to a more precise multidimensional construct with four interrelated domains.

Keywords:

student evaluation of teaching, scale validation, higher education, item response theory

1 Introduction

Student evaluation of teaching (SET) ratings is a generalized practice in almost every institution of higher education (Huybers, 2014; Richardson 2005; Zabaleta, 2007). Regarded as eliciting perceived performance feedback on a range of teacher-and/or course-related aspects, in most universities, SET is used for formative purposes (e.g. feedback for the improvement of instruction) as well as for administrative decision-making (e.g., about recruitment, career progress, and economic incentives) (Linse, 2017).

Considering the educational dynamic context as one of the most relevant factors in any teaching-learning process, the purpose of the present study was to develop a measure of SET from a large public institution in Ecuador. The access to institution-wide student survey data may reveal possible structural or psychological barriers, considering the underlying relations among beliefs, knowledge, and actions involved in the field.

1.1 Theoretical perspectives on SET

The analysis of the teaching process has been raised as an important concern in the last decades which has been related to the diversity of models of teaching. Initially, teacher competency moved from on a series of actions and behaviors (Boice, 1991) to a more complex cognitive activity (Leinhardt & Greeno, 1986) and affections (Berlinger, 1986).

Given these diverse structures, the initial frameworks of effective teaching started with the commonly named theories of expertise, in which the most valuable information was reported through indicators like depth of problem representation, knowledge organization and structure, efficiency of procedures, and metacognitive skills for learning, among others (Glaser, Lesgold, & Lajoie, 1984). However, particularly in higher education, beliefs or views about teaching have also been considered as important constructs in effective teaching (Larsson, 1986). Considering beliefs as "implicit assumptions about students, learning classrooms, and the subject matter" (Pajares, 1992; Pratt, 1997), this framework moves toward a developmental and integrated view of teaching and the learning process as a necessary link to an effective SET.

Within a more dynamic perspective, the theory of transformative learning (Mezirow, 1991) posits an important distinction between individuals' previous beliefs of teaching and what they actually do when they teach. In those situations, the identification of contextual barriers is crucial to determine, as well as the identification of teachers' lack of effective strategies to implement positive assumptions (Schön, 1983). The change towards effective teaching lies in an effective engagement process with support and application of strategies in positive learning conditions.

In this setting, the authors of the present study relied on "the new model of teaching" provided by Saroyan and Amundsen (2001). This ecological framework considers three main elements associated with SET: conceptions or beliefs, knowledge and actions. Moreover, the key element for dynamic formation of the concepts are based on the analyses of the contextual influences. In this latter construct, it includes all external factors which may influence teaching tasks, such as the culture of the university, faculty or department, or instructors' teaching assignments (Saroyan & Amundsen, 2001).

Noting this ecological perspective, in Latin-American higher education institutions, special efforts have been taken to be considered as relevant actors of social development (Arocena & Sutz, 2005). During the twentieth century, important social movements triggered the so- called University Reform Movement (URM) (Ribeiro, 1971), allowing the inclusion of social policies in higher education, in spite of political or military controversies through the past decades. It can be said that this spread of democracy in the higher education system has met the goals of stronger teaching and research standards.

The Council of Ecuadorian Higher Education established the obligatory nature of the evaluation of the teaching in the higher education institutions, both for its organization and promotion, in the Career and Ladder Regulations of the Professor and Researcher of the Higher Education System (CES, 2017). The assessment of the performance of university teachers is an essential component that allows a professor to enter as an Assistant Professor or Associate Professor. The requirements include a score of at least 75% in the performance evaluation during the last two academic periods. Additionally, according to Article 96 of the regulations, members of the academic staff will be dismissed if they have obtained: 1) two consecutive integral evaluation of performance scores of less than 60%; and 2) four integral evaluations of performance scores of less than 60% during their careers. In addition, it establishes that the main titular teachers will be promoted to the next higher level if they comply with other requirements such as having obtained at least 80% on their performance evaluation scores in the last two academic periods (Consejo de Educación Superior [CES], 2017).

1.2 Connecting Educational and Learning Capitals through SET

Recently, synthetic perceptions of achievement levels have been proposed as an effective strategy to overcome analytic strategies (Veas et al., 2018). From the giftedness field, the transformation of talents, gifts, or abilities into achievements is still considered a linear sum of independent variables. This basic assumption is also virtually expressed in graphic representations of models when neatly separated boxes of variables are listed after bullet points. However, usually no information is given about the exact nature of the interplay and the involved processes (Heller et al., 2005).

A synthetic perspective of achievement can be proposed from an extension of the actiotope model of giftedness (Ziegler et al., 2017; Ziegler & Stoeger, 2017) to all achievement levels, as every student constitutes an actiotope with specific resources. Concretely, the influx of exogenous resources from the environment is of particular importance. In this context, Ziegler and Baker (2013) referred environment resources as educational capital, whereas internal resources are considered as learning capital -resources that can be used to promote learning). Within the SET perspective, student's criteria consist on a cultural educational capital, which includes thinking patterns which can facilitate – or hinder- the attainment of learning and educational goals (Ziegler & Baker, 2013, p. 28). For this reason, this resource directly affects to the teachers' view of their own professional development, which can be composed by telic learning capital and actional learning capital (Ziegler et al., 2017; Ziegler et al., 2019). Indeed, evaluation of teaching is effective when educational systems may pro-

mote changes on teachers' goals to activate professional actions, and therefore optimal results.

1.3 Review of SET measures

The instruments normally used to measure students' evaluation of their teachers, programs, and students' satisfaction with their instruction, are the standard rating scales. However, research on SET ratings have not yet provided a clear answer about some questions of their validity (Hornstein, 2017; Marsh, 2007a, b; Spooren, Brockx, & Mortelmans, 2013; Uttl et al., 2017). Interestingly, many of the evaluation instruments have been constructed and validated within the institution itself, and the results of this validation have not always been published nor they have even been tested for their psychometric quality (Richardson, 2005).

Several well-designed and validated SET instruments are available (Spooren et al., 2013). One of these instruments more widely used is the one published by Marsh (1982), and Marsh et al (2009), the Student's Evaluations of Educational Quality (SEEQ). This scale is composed of 35 statements which to each the students respond, using a five-point scale. The scale evaluates nine aspects of teaching: learning value, enthusiasm, organization, group interaction, individual rapport, breadth of coverage, exams and grading, assignments, workload or difficulty. These dimensions have been reproduced in the Confirmatory Factor Analyses, using large samples in different countries, different teacher status, and disciplines (Marsh, 1987; 2007a). Other more recently developed assessment instruments are the Students' Evaluation of Teaching Effectiveness Rating Scale (SETERS) by Toland and De Ayala (2005), the Teaching Proficiency Item Pool (Barnes et al., 2008), the SET37 inventory by Mortelmans and Spooren (2009) and the Teaching Behavior Checklist (Keeley, Furr, & Buskist, 2010).

However, there is no consensus on the number and type of dimensions. This lack of consensus (Apodaca & Grad, 2005; Spooren et al., 2013) is due to conceptual problems related to the lack of a common theoretical framework about what effective teaching is, and methodological problems concerning the measurement of dimensions, as a data-driven process in which different post hoc analytic techniques are used. It seems necessary to use the most common dimensions that are associated with greater teaching effectiveness.

The question concerning construct validity that arises in relation to SEEQ and other instruments is about its unidimensional or multidimensional structure. Researchers agree that there are several dimensions, but it is not clear whether they can be subsumed into a single global dimension. Marsh (1987, 1991a, b, 2007) considers that, although the dimensions of the SEEQ were correlated with each other, they are not represented by a general higher-order factor.

Against the assertion of Abrami, d'Apollonia, and Rosenfield (1997) that the SEEQ dimensions were subsumed by a single construct called "general instructional skill"; based on the evidence from the results of Hierarchical Confirmatory Factor Analysis (HCFA), Marsh (1991b, 2007a) concluded that support for a multidimensional view is strong. Conversely, using HCFA, Cheung (2000) found Evidence of a Single Second-Order Factor in Student Ratings of Teaching. Marsh et al (2009) defended a multidimensional structure of the students' evaluations of university teaching (SETs), on the basis of which measures can be ob-

tained: both of the specific dimensions and of a general factor of the quality of teaching.

The test-retest reliability of students' evaluations is high, even when there is a long period of time between evaluations (Richardson, 2005). There is a high correlation between the scores of students taking different subject-matter taught by the same teacher, but a low relationship between the evaluations given by students taking the same subject-matter taught by different teachers. This suggests that students' evaluations are a function of the teacher's appeal rather than the subject matter (Marsh, 2007). Moreover, evaluations of the same teachers given by successive cohorts of students are highly correlated (Marsh, 2007b; Marsh & Hocevar, 1991; Marsh et al., 2009).

The inter-rater reliability of individuals and the average ratings given by groups of students was commonly high (Marsh and Roche 1997; Richardson, 2005); although the research by Feistauera and Richterb (2017) suggested that student evaluations of teaching can be reliable assessments of the course and the teacher when aggregated evaluations based on a sufficient number of students are used; however, the inter-rater reliability of student evaluations of teaching varied between different measures and course types (seminar and lecture).

The present study was carried out in a different context to most of the previous works (Clayson, 2009). Concretely, it measures the student evaluation of teaching in a higher education institution, the National Polytechnic School of Ecuador, where students are enrolled in technical subjects such as engineering, architecture and biotechnology. Although there were no records in the beginning of teacher evaluations in higher education in Ecuador, this has been a widespread practice in Ecuadorian higher education institutions since the early 1980s (Pareja, 1986).

The student evaluation of teaching instrument used in the National Polytechnic School is the "Cuestionario de Evaluación de la Enseñanza del Profesor de la Escuela Politécnica Nacional del Ecuador" [*Teacher Evaluation Questionnaire of the National Polytechnic School, TEQNS*]. The evolution of the questionnaire consisted of the proposal of several effective teaching criteria, from which a set of items were developed by a teaching committee, which was part of the management team of the *National Polytechnic School.* The aspects to be evaluated and the specific items that make up the questionnaire are approved each academic year by the management team of the National Polytechnic School.

To determine the dimensions of teaching competence and more specific aspects to evaluate, both conceptual and applied aspects were taken into account. Based on these two approaches, Apodaca and Grad (2005) proposed five dimensions: 1. planning and preparation, 2. communication skills 3. interaction with students 4. didactics and methodological resources, and 5. assessment.

From the perspective of academic staff training, Newble and Cannon (1995) proposed 'organization', 'instruction', 'evaluation', 'relationships' and 'subject mastery' as the most important aspects of teaching. Therefore, the items are grouped theoretically into the following four factors: 1. *Planning, mastery and clarity in the explanation of the subject matter*, that incorporates the knowledge of subject, clarity and understanding, including sensitivity to and concern with class level and progress, structure, planning, preparation and organization of the course (i.e. *The teacher appropriately*

selected the class activities, according to the objectives). 2. *Methodology and resources,* referring to the use of appropriate and varied teaching methods and materials (i.e. *The teacher conveniently used different teaching methods*). 3. *Evaluation,* understood as the use of objective and impartial methods, related to teaching, and useful to reorient student learning (i.e. *The teacher evaluated fairly and impartially*). 4. *Teacher-student relationship*, referring to concern and respect for students, friendliness of the teacher, rapport, openness to opinions of students, encouragement of the student initiatives, availability, and helpfulness (i.e. *The teacher has been given suggestions that he/she accepted openly*).

Although the number and dimensions of effective teaching have remained an open question (Spooren et al., 2013), these four dimension are present in the most of SET rating scales literature, and they are aspects related to the teaching effectiveness (Apodaca and Grad, 2005; Cohen, 1981; Feldman, 1989; Huybers, 2014; Richardson, 2005; Spooren, et al., 2013).

Content and face validity are taken mainly into account in the development of the questionnaires. However, the empirical validation is minimal and is limited to the descriptive and discriminatory analysis of the items individually considered. It is lacking a complete process of construct and criterion validity, just as the estimation of the reliability of the scale and/or the subscales that make up these instruments. In this regard, item response theory (IRT) has the main advantage of focusing on the quality of items in measuring underlying constructs (Van der Linden, 2017). IRT models give researchers more confidence in applying the scale in wider contexts (Wright & Masters, 1982). Differing from classical test theory, which considers that an observed test score is composed by a true score and a random error component, IRT considers that the probability of a person's expected response to an item is a mathematical function of that person's ability and one or more parameters characterizing the item (Bond & Fox, 2015, p. 363).

The Rasch model (Rasch, 1960) is the most well-known among unidimensional item response theory (UIRT) models, providing a method based on the calibration of ordinal data from a shared measurement scale and enabling one to test conditions such as dimensionality, linearity and local independence. Calibration is the procedure used to estimate personal latent traits or item difficulty by converting raw score odds to logits on an IRT measurement scale (Bond & Fox, 2015). Moreover, as the TEQNS has different factors, a multidimensional item response theory (MIRT) is a better technique to simultaneously calibrate all subscales and increase the measurement precision by taking into account the correlation between subscales (Adams, Wilson, & Wang, 1997).

The aim of the present study was to examine the internal validity of the TEQNS, applying both unidimensional item response theory (UIRT) and multidimensional item response theory (MIRT) models to examine and compare different theoretical internal structures of the instrument. Concretely, three research questions were considered:

Research question 1 (RQ1): Do all the TEQNS items represent a unique construct?

Research question 2 (RQ2): Do each of the different factors of the TEQNS represent a single unidimensional construct?

Research question 3 (RQ3): Are the factors of the TEQNS interrelated?

2 Method

2.1 Participants

The sample consisted of 6110 students of the National Polytechnic School from Ecuador, who rated the teaching of their 310 teachers, who composed a varied sample of age, category, and teaching experience. These students were enrolled in 8 different Faculties and Schools, in 28 different degree programs, and attended 358 different classes. In the population, 68.3% of the students were male and 31.7% female. The distribution of students and percentages per academic department and undergraduate degrees can be seen in Table 1. The higher percentage of male students was representative of the population of students of polytechnic studies. The average age was 22.6 years old (SD = 3.2). These students rated the faculty's teaching during the 2016/17 academic year.

Table 1 Frequencies and gender distribution from the TEQNS dataset by faculty and undergraduate degree

Faculty	Undergraduate degree	Frequency	Gender	
			M	F
Basic Sciences	Mathematics	33	20	13
	Physics	106	80	26
	Basic Sciences	90	62	28
Administration Sciences	Bussiness Studies	315	180	135
	Economic and Finantial Studies	232	150	82
	Social Sciences	282	145	137
Civil and Environmental Engineering	Civil Engineering	342	222	120
	Environmental Engineering	264	166	98
	Water Technology	168	88	80
Electronic Engineering	Electric Engineering	281	192	89
	Electronic and control Engineering	462	351	111
	Information network Engineering	225	157	68
	Telecommunitations Engineering	359	244	115
	Mathematical Engineering	397	313	84
	Electromechanical Technology	172	113	59
	Electronical Technology and Telecommunications	350	255	95
Geology and Petroleum	Geological Engineering	111	81	30
	Petroleum Engineering	202	177	25
Mechanical Engineering	Mechanical Engineering	710	532	178
Chemical engineering and Agribusiness	Chemical Engineering	283	187	96
	Agribunsiness Engineering	170	102	68
Computer Systems Engineering	Computer Systems and Computing Engineering	439	285	154
	Computer Technology	117	71	46
Total		6110	4173	1937

Note. M= Male. F= Female

2.2 Measures

Students' evaluations of teaching ratings were obtained from the "Cuestionario de Evaluación de la Enseñanza del Profesor de la Escuela Politécnica Nacional del Ecuador", [*Teacher Evaluation Questionnaire of the National Polytechnic School*], approved by the teaching staff for the 2016-17 academic year.

The items are grouped theoretically into four factors: 1. Planning, mastery and clarity in the explanation of the subject-matter (items 1-9), (i.e.: *The teacher conveniently expressed the objectives and themes, indicating their relationship with the professional training of the studies taken*). 2. Methodology and resources (items 10-15), (i.e.: *The teacher explained didactic material apart from the textbook and made it understandable*). 3. Evaluation (items 16-23) (i.e.: *The evaluation events are related to the teaching given*). 4. Teacher-student relationship (items 24-32 items) (i.e.: *The teacher created a climate of trust and work in class*). Students responded to these items on a 5-point rating scale ranging from 1 (do not agree at all) to 5 (totally agree). The full scale, with the items grouped into the four theoretical dimensions are included in the Annex, in original Spanish version and English translation.

2.3 Procedure

The data collection was made from the existing computer records in the administration of the National Polytechnic School of the Ecuador and, access was granted with permission of the Vice Chancellor for Academic Affairs of the Institution. The data provided by the institution were anonymous, with only one identification code for each student. Students' age, and gender, as well as teachers' age, gender, and experience were collected from administrative records.

The application of the scale of evaluation of teaching by the student was carried out towards the end of the semester, before they knew their final grades. All the teachers were evaluated by the students in a similar period of time. All the students had to evaluate the teachers to be able to access their final grades. The student evaluation of teaching was made through an electronic platform, in which the data were recorded.

The impact of faculty procedures on response rates of student evaluations of teaching has been studied by several authors, as opposed to special electronic evaluations. So, Young, Joines, Standish, and Gallagher (2019) found that response rates were substantially higher when faculty provided in-class time for students to complete student evaluations of teaching compared to the electronic form from the administration. However, there are studies designed to analyze this question that do not find differences between the evaluations with electronic questionnaires and those with paper and pencil; additionally, this is true when a more representative sample responds instead of a smaller and biased sample (Nowell, Gale & Kerkvliet, 2014).

Once, the response rate in electronic administration was lower than with paper-and-pencil questionnaires, so the procedure followed in this case consisted of forcing all the students to answer the evaluation survey in order to access their final grades. This procedure has proved useful and valid in some higher education institutions (Leung, & Kember, 2005; Nair, & Adams, 2009).

2.4 Data analysis

Two indicators, item weighted fit and "expected a posteriori/plausible value (EAP/PV)" reliability, were examined. Item weighted fit was selected as it indicates how well the item parameters of a measure fit the empirical dataset. Concretely, infit and outfit statistics were used to check the quality of the instrument. These indexes are the mean value of the squared residuals. Therefore, the larger the squared residual, the larger misfit between data and model.

The infit statistic is an information-weighted sum, so this variance is larger for well-targeted observations and smaller for extreme observations (Bond & Fox, 2015). d values of outfit and infit mean squares can range from 0 to positive infinitive. Values below 1 indicate a higher than expected fit of the model, whereas values greater than 1 indicate poor fit of the model. An infit/outfit range between 0.75 and 1.33 can be considered as acceptable values (Wu, Adams, & Wilson, 1998).

With respect to EAP/PV reliability, it is the ratio of modeled variance to observe variance. The interpretation of this parameter is similar than Cronbach's alpha, such that a reliability .70 is considered to be the minimum standard, while .80 is recommended for screening purposes (Salvia, Ysseldyke, & Bolt, 2013).

The analyses were involved in three phases: In the first phase, the UIRT of model 1 was examined, which is based on the assumption that all the items load on a unique construct, namely "student evaluation of teaching". Item weighted fit and EAP/PV reliability, and principal component analyses were the indicators used to examine this model. Winsteps version 4.5 statistical software (Linacre, 2019b) was used to check whether the items in each subscale satisfy the unidimensionality assumption. Unidimensionality requires that the measurement should target one attribute or dimension at one time (Bond & Fox, 2015). According to Linacre (2019a), an eigenvalue less than 2.0 of the first contrast indicated the residuals are not relevant enough to disturb the measurement quality. An eigenvalue more than 2.0 implies that there is probably another dimension in the measurement instrument.

In the second phase, it examined the quality measurement of the items within each domain independently, which a total of 4 factors (e.g., Model 2-PMC; Model 2-MR; Model 2-E, and Model 2-TR). Item weighted fit and EAP/PV reliability were the indicators used to examine the goodness of fit.

The third phase consisted of examining Model 3, which represented the hypothetical model constructed for the TEQNS. To evaluate it, Model 3 was calibrated to evaluate the item weighted fit, EAP/PV reliability, and the correlation between latent traits by domain. Results from Model 3 were first compared with results from Model 2 to see whether the EAP/PV reliability for each factor was improved. Results from Model 1 were also compared with Model 3 in order to assess which model shows the best underlying measurement description of both items and factors. To this end, deviance (-2log likelihood) and Akaike information criterion (AIC) were employed (Wu, Adams, Wilson, & Haldane, 2007).

The internal structure of the TEQNS was tested through various IRT models using Conquest version 2.0 (Wu et al., 2007). Each of the three models are summarized in Table 2. All models are based on the rationale of the Rasch Rating Scale Model (RRSM). Maximum likelihood estimation method

was employed for the parameters of the model. The Monte Carlo method was used for calibration in all the models to ensure index comparison.

3 Results

With respect to the first model, the TEQNS was considered to measure a unidimensional construct, namely, student evaluation of teaching as an overall composed dimension. The weighted fit was acceptable in all items with the exception of item 11 (see Table 3). The EAP/PV reliability was 0.91, considered as an excellent value. For this reason, a complement analysis of unidimensionality was used. A principal component analysis of the residual scores (Linacre, 1998; Wright, 1996) showed that the eigenvalue of the first contrast for the whole scale was 3.61; the second contrast was 2.64 and the third contrast was 2.48. Therefore, the instrument could be considered as multidimensional.

Table 2 Summary of IRT models

	Model 1	Model 2	Model 3
IRT model	UIRT	UIRT	Between-item MIRT RRSM
Theory	All items represent a single construct	Items within each factor represent independent constructs	All items represent four interrelated domains.
Number of items	32	F1= 9 items; F2 = 6 items; F3 = 8 items; F4 = 9 items	32 items (9 in F1; 6 in F2; 8 in F3; 9 in F4).
Answering research question	1	2	3

Note. Model 2 represented each of the four domains, named as Model 2-PMC (Planning, mastery and clarity in the explanation of the subject), Model 2-MR (Methodology and resources), Model 2-E (Evaluation), and Model 2-TR (Teacher-student relationship). IRT = Item Response Theory; UIRT = unidimensional item response theory; RRSM = Rasch Rating Scale Model; MIRT = multidimensional item response theory.

Table 3 Fit statistics information for models

	Weighted fit				Misfit items	
Model	Minimum	Maximum	M	SD	Overfit	Underfit
Model 1	0.81	1.18	1.00	0.18	item 11	-
Model 2-PMC	1.00	1.34	1,13	0,11	item 8	-
Model 2-MR	0.92	1.44	1.06	0.20	item 11	-
Model 2-E	0.93	1.45	1.08	0.18	item 16	-
Model 2-TR	0.94	1.30	1.08	0.11		-
Model 3	0.85	1.68	0.99	0.15	item 11	-

Note. "-" indicates that no underfit values were found; PMC = Planning, mastery and clarity in the explanation of the subject; MR = Methodology and resources; E = Evaluation; TR = Teacher-student relationship

In Model 2, the four factors of the TEQNS were calibrated independently within four unidimensional structures. Results indicated that only Model 2-TR had adequate fit statistics, whereas both in Model 2-PMC and Model 2-MR had one misfit item each. In all the models, the EAP/PV values were adequate.

Model 3 was calibrated within MIRT, in which the four domains were inter-correlated. Results from the weighted fit statistics showed adequate model fit for most of the items, as presented in Table 3. Similar to Model 1 and Model 2-MR, item 11 presented extreme fit values. EAP/PV reliabilities showed excellent values, all of the higher than .90 as can be seen in Table 4. Correlations between domains are showed in Table 5. As can be observed, all the correlations ranges were high, and even the correlation between Evaluation and Teacher-student relationship was extremely high (r = .93). When comparing Models 1 and Model 3, it seems that the last set showed better deviance and AIC values (see Table 6). The deviance difference can be considered as statistically significant.

Given that Model 3 showed the best psychometric properties, item 11 was removed due to its misfit value, which was consistent in Model 1 and Model 2-MR. A re-analysis of Model 3 without this item showed appropriate weighted fit values for all the items, ranging from 0.91 to 1.18.

Table 4 EAP/PV Reliability Summarized by Models.

Model	PMC	MR	E	TR	Total
Model 1	-	-	-	-	0.91
Model 2-PMC	0.89	-	-	-	
Model 2-MR	-	0.91			
Model 2-E	-	-	0.99	-	-
Model 2-TR	-	-	-	0.91	-
Model 3	0.95	0.94	0.98	0.95	-

Note. "-"indicates the EAP/PV reliability was not calculated by the specific model. EAP/PV = expected a posteriori/plausible value; PMC = Planning, mastery and clarity in the explanation of the subject; MR = Methodology and resources; E = Evaluation; TR = Teacher-student relationship.

Table 5 Correlation between subdomains estimated for Model 3

Domains	PCM	MR	E
PCM			
MR	0.87		
E	0.90	0.90	
TR	0.87	0.88	0.93

Note. PCM = Planning, mastery and clarity in the explanation of the subject; MR = Methodology and resources; E = Evaluation; TR = Teacher-student relationship.

Table 6 Comparison of Model fit statistics

Model/Comparison		Deviance	AIC	N of parameters
Model				
	Model 1 (unidimensional)	125004.41	125076.41	36
	Model 3(multidimensional)	118823.56	118913.56	45
Model Comparison				
	Model 1-Model 3	6180.85*	6162.85	9

Note. AIC = Akaike information criterion.

*p<.001 when deviance is greater than the critical value of chi-square distribution (χ^2 = 27.87, df = 9).

4 Discussion

Student evaluation of teaching is considered as one of the most important procedures in different kinds of institutions, which include a variety of rating procedures with a lack of a clear consensus (Hornstein, 2017). Based on an extension of the actiotope model of giftedness, which considers the relevance of educational and learning capitals on achievement, this study applied both UIRT and MIRT models to analyze the underlying structure of the TEQNS on a large sample of university students from the National Polytechnic School of Ecuador. This instrument was developed to address methodological problems that exist in the current teacher- student evaluation measures, mainly related to construct validity.

In the first place, in response to RQ1, a calibration of unidimensional measure of SET, showed adequate fit statistics in most of the items, as well as a high total reliability score. This suggested that the total score of the TEQNS may reflect a composite score of the construct. The results also addressed the possibility that the instrument could measure four unidimensional and independent domains (RQ2). In this case, the results showed a higher number of misfit items, and adequate reliability values. In addition, the third question (RQ3) was based on the construction of four inter-correlated domains through a MIRT model, which was supported by adequate reliability and correlation between domains.

When comparing the deviance and AIC of Model 3 with the deviance and AIC of Model 1, it showed that Model 3 has a better description of the data. This indicated that the nature of the internal structure in TES was multidimensional and involved multiple distinct domains. Furthermore, when comparing Model 3 and Model 2, it is recognized that Model 3 has better EAP/PV reliability estimates in three out of the four domains.

It is important to mention that item 11 (*The teacher organized didactic experiences such as visits, excursions, projects, discussions*) showed a constant misfit in all the models where it was included (Model 1, Model 2-PMC, and Model 3). This is a concern affecting the underlying construct. It is possible that the content of this item may not have a direct relation with the student's conception of methodological and didactic resources, but with emotional or social group experiences that exceeds the teach-

ers' organizational features. For this reason, a new calibration of Model 3 without this item was considered, detecting a better distribution of item parameters, with weighted fit statistics rating from 0.91 to 1.18 and no misfit items in any dimension.

This is the first study that implements MIRT model as an effective measure of SET in higher education. Furthermore, given the need to have valid and reliable tools for the assessment of SET in Latin America, the construct validation of the TEQNS has a strong role to play in the assessment of teachers in large public universities.

The current study provides a dynamic perspective on domains of SET, adding important information regarding not only general model fit indexes (as happens with traditional Exploratory or Confirmatory Factor Analysis), but also specific parameters which measure the adequacy of each item as an effective indicator of the underlying construct. IRT can be considered as an effective modeling approach to ensure SET valid measures in the dynamic perspective of educational and learning capitals (Ziegler & Baker, 2013; Ziegler et al., 2017). In line with the characterizing features of a synthetic research strategy, the comparison of endorsability of each item is no longer dependent on a specific sample, and the comparison of the student's assessment of teaching is no longer dependent on specific items (Hambleton & Jones, 1993). At the same time, Rasch models can examine how participants understand different response options and provide additional evidence on the internal construct for SET subscales.

4.1 Limitations and future directions

Although a large dataset was employed for psychometric validation of the TEQNS in one of the most important public institutions in Ecuador, bias analysis was not included in this study. Its results were limited to this institution and this type of technical studies.

A special source of bias in SET studies is response bias, a respondent's consistent answering pattern irrespective of the questions presented. Several studies found both acquiescent responding and extreme responding to be consistent traits within individual students; however, when these were statistically controlled, although their effects were reduced, the relationship remained the same (Huybers, 2014; Richardson, 2012).

As a more general sense, bias is present when a known characteristic of students systematically affects their ratings of teachers. The gender of the students is an example of this possible bias in student evaluation of teaching. Previous studies have found that female students on average tend to give significantly higher SET ratings than their male peers (Badri, Abdulla, Kamali, & Dodeen, 2006; Basow, Phelan, & Capotosto, 2006; Boring, 2015; Centra, & Gaubatz, 2000; Darby, 2006).

Another source of bias is the discipline. If the evaluation of teaching is situational and is affected for academic disciplines, being higher in studies in the field of education and the liberal arts, and not as high in other areas, such as business and engineering (Clayson, 2009) it seems necessary to carry out new studies in areas different from the previous ones, as the technical areas where there are less studies on the subject.

Two steps in the development of this measure should be considered: first, personal and psychological variables should be analyzed, including students' demographic characteristics (e. g. geographical regions, race) to check measurement precision of the instrument. Second, further analyses are required to detect possible causes of SET differences between sample subgroups and efficient vs non-efficient teachers' features associated with SET. Such work would further both the reliability and validity of the current measurement using diverse populations and ecological contexts.

5 References

Abrami, C.P., d'Apollonia, S., & Rosenfield, S. (1997). The dimensionality of student ratings of instruction: what we know and what we do not. In R. E. Perry and J.C. Smart (Eds.), Effective teaching in higher education: Research and practice (pp. 321–367). New York: Agathon Press.

Adams, R. J., Wilson, M., & Wang, W. C. (1997). The multidimensional random coefficient multinomial logit model. *Applied Psychological Measurement, 21*, 1-23. https://doi.org/10.1177/0146621697211001

Apodaca, P., & Grad, H. (2005). The dimensionality of student ratings of teaching: Integration of uni- and multidimensional models. *Studies in Higher Education*, *30*, 723–748. https://doi.org/10.1080/03075070500340101

Arocena, R., & Sutz, J. (2001). Changing knowledge production and Latin American universities. *Research Policy, 30*(8), 1221-1234.

Badri, M. A., Abdulla, M., Kamali, M. A., & Dodeen, H. (2006). Identifying potential biasing variables in student evaluation of teaching in a newly accredited business program in the UAE. *International Journal of Educational Management, 20*(1), 43-59.

Barnes, D., Engelland, B., Matherne, C., Martin, W., Orgeron, C., & Ring, J.(2008). Developing a psychometrically sound measure of collegiate teaching proficiency. *College Student Journal*, *42*, 199-213.

Barrett, P. (2012). *Gower program help file*. Auckland, New Zealand: Advanced Projectes R&D Ltd.

Basow, S. A., Phelan, J. E., & Capotosto, L. (2006). Gender patterns in college students' choices of their best and worst professors. *Psychology of Women Quarterly*, *30*(1), 25-35.

Boice, R. (1991). New faculty as teachers. *Journal of Higher Education, 62*(2), 150-173.

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). New York, NY: Routledge.

Centra, J.A., and Gaubatz, N.B. (2000). Is There Gender Bias in Student Evaluations of Teaching? *The Journal of Higher Education*, *71*(1), 17-33. https://doi.org/10.2307/2649280

Cheung, D. (2000). Evidence of a single second-order factor in student ratings of teaching effectiveness. *Structural Equation Modeling: A Multidisciplinary Journal*, *7*(3), 442-460, DOI: 10.1207/S15328007SEM0703_5

Cho, E. (2016). Making reliability reliable: A systematic approach to reliability coefficients. *Organizational Research Methods*, *19*(4), 651-682. https://doi.org/10.1177/1094428116656239

Clayson, D. E. (2009). Student evaluations of teaching: Are they related to what students learn? A meta-analysis and review of the literature. *Journal of Marketing Education*, *31*, 16–30. https://doi:10.1177/0273475308324086

Consejo de Educación Superior (CES) (2017). Reglamento de Carrera y Escalafón del Profesor e Investigador del Sistema de Educación Superior. [Career and Ladder Regulations of the Professor and Researcher of the Higher Education System]. Retrieved on the 20 April 2019, from: https://procuraduria.utpl.edu.ec/sitios/documentos/NormativasPublicas/Reglamento%20de%20Carrera%20y%20Escalaf%C3%B3n%20del%20Profesor%20e%20Investigador%20del%20Sistema%20de%20Educaci%C3%B3n%20Superior%202018.pdf

Darby, J. A. (2006). Evaluating courses: An examination of the impact of student gender. *Educational Studies, 32*(2), 187-199. https://doi.org/10.1080/03055690600631093

Feistauera, D., & Richterb, T. (2017). How reliable are students' evaluations of teaching quality? A variance components approach. *Assessment & Evaluation in Higher Education*, *42*(8), 1263–1279. https://doi.org/10.1080/02602938.2016.1261083

Glaser, R., Lesgold, A, & Lajoie, S. P. (1988). Toward a cognitive theory for the measurement of achievement. In R. Ronning, J. Glover, J. S., Conoley, & J. C. Wittrock (Eds.). *The influence of cognitive psychology on testing*, vol. 3. Hillsdale, NJ: Erlbaum.

Heller, K. A., Perleth, C., & Lim, T. K. (2005). The Munich model of giftedness designed to identify and promote gifted students. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (2nd ed., pp. 147-170). Cambridge, UK: Cambridge University Press.

Hornstein, H. A. (2017). Student evaluations of teaching are an inadequate assessment tool for evaluating faculty performance. *Cogent Education, 4*, 1. http://dx.doi.org/10.1080/2331186X.2017.1304016

Huybers, T. (2014) Student evaluation of teaching: the use of best–worst scaling, *Assessment & Evaluation in Higher Education, 39*(4), 496-513. https://doi.org.10.1080/02602938.2013.851782

Keeley, J., Furr, R.M., & Buskist, W. (2010). Differentiating psychology students' perceptions of teachers using the Teacher Behavior Checklist. *Teaching of Psychology*, *37*, 16-20. https://doi.org.10.1080/00982890342682

Larsson, S. (1986). Learning from experience: teachers' conceptions of changes in their professional practice. *Journal of Curriculum Studies, 19*(1), 37-43.

Leinhardt, G., & Greeno, J. G. (1986). The cognitive skill of teaching. *Journal of Educational Psychology, 78*(2), 75-95.

Leung, D.Y.P., & Kember, D. (2005). Comparability of data gathered from evaluation questionnaires on paper and through the Internet. *Research in Higher Education*, *46*(5), 571-591. https://doi.org/10.1007/s11162-005-3365-3

Linacre, J. M. (1998). Structure in Rasch residuals: Why principal component analysis? *Rasch Measurement Transactions, 12*, 636.

Linacre, J. M. (2019a). *A user's guide to Winsteps & Ministeps Rasch-Model computer programs*. Available at http://www.winsteps.com

Linacre, J. M (2019b). *Winsteps* (version 4.4.7.) [computer software]. Chicago: MESA.

Linse, A.R. (2017). Interpreting and using student rating data: Guidance for faculty serving as administrators and on evaluation committees. *Studies in Educational Evaluation*, *54*, 94-106. https://doi.org/10.1016/j.stueduc.2016.12.004

Marsh, H. W. (1982). SEEQ: a reliable, valid and useful instrument for collecting students' evaluations of university teaching. *British Journal of Educational Psychology*, *52*, 77-95.

Marsh, H. W. (1987). Student's evaluations of university teaching: Research findings, methodological issues, and directions for further research. *International Journal of Educational Research*, *11*, 253–388. https://doi.org/10.1016/0883-0355(87)90001-2

Marsh, H. W. (1991a). A multidimensional perspective on student's evaluations of teaching effectiveness: Reply to Abrami and d'Apollonia (1991. *Journal of Educational Psychology*, *83*,(3), 416–421. https://doi.org/10.1037/0022-0663.83.2.285

Marsh, H. W. (1991b). Multidimensional students' evaluations of teaching effectiveness: A test of alternative higher-order structures. *Journal of Educational Psychology*, *83*, 285–296. https://doi.org/10.1037/0022-0663.83.2.285

Marsh, H. W. (2001). Distinguishing between good (useful) and bad workloads on students' evaluations of teaching. *American Educational Research Journal*, *38*, 183–212. https://doi.org/10.3102/00028312038001183

Marsh, H. W. (2007a). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In R. P. Perry & J. C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence based perspective* (pp. 319–383). New York: Springer.

Marsh, H. W. (2007b). Do university teachers become more effective with experience? A multilevel growth model of students' evaluation of teaching over 13 years. *Journal of Educational Psychology*, *99*, 775–790. https://doi.org/10.1037/0022-0663.99.4.775

Marsh, H.H., Hau, K.T., Balla, J.R., & Grayson, D. (1998). Is more ever too much? The number of indicators per factor in confirmatory factor analysis. *Multivariate Behavioral Research*, *33*, 181-230.

Marsh, H.W., & Hocevar, D. (1991). Students' evaluations of teaching effectiveness: the stability of mean ratings of the same teachers over a 13-year-period. *Teaching and Teacher Education*, *7*, 303-314.

Marsh, H. W., Martin, A. J., & Jackson, S. E. (2010). Introducing a short version of the physical self description questionnaire: new strategies, short-form evaluative criteria, and applications of factor analyses. *Journal of sport & exercise psychology*, *32*(4), 438-482. https://doi.org/10.1123/jsep.32.4.438

Marsh, H. W., Muthèn, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A. J. S., & Trautwein, U. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling*, *16*, 439–476. https://doi.org/10.1080/10705510903008220

McNeish, D. M. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, *23*(3), 412-433. https://doi.org/10.1037/met0000144

Mezirow, J. (1991). *Transformative dimensions of adult learning*. San Francisco: Jossey-Bass.

Morin, A. J. S., Katrin Arens, A., & Marsh, H. W. (2016). A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Structural Equation Modeling*, *23*(1), 116-139. https://doi.org/10.1080/10705511.2014.961800

Mortelmans, D., & Spooren, P. (2009). A revalidation of the SET37-questionnaire for student evaluations of teaching. *Educational Studies*, *35*, 547-552. https://doi.org/10.1080/03055690902880299

Muthen, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology*, *25*, 267-316. https://doi.org/10.2307/271070

Muthén, L. K., & Muthén, B. O. (2018). *Mplus user's guide* (8.ª ed.). Los Angeles, CA: Muthén & Muthén.

Nair, C.S., & Adams, P. (2009) Survey Platform: A Factor Influencing Online Survey Delivery and Response Rate, *Quality in Higher Education*, *15*(3), 291-296. https://doi.org/10.1080/13538320903399091

Nagy, M. S. (2002). Using a single-item approach to measure facet job satisfaction. *Journal of Occupational and Organizational Psychology*, *75*(1), 77–86. http://dx.doi.org/10.1348/096317902167658

Newble, D. & Cannon, R. (1995). *A handbook for teachers in universities & colleges: a guide to improving teaching methods*. London: Kogan Page.

Nowell, C., Gale, L.R., Kerkvliet, J. (2014). Non-response bias in student evaluations of teaching. *International Review of Economic Education*, *17*, 30-38.

Pajares, M. F. (1992). Teachers' beliefs and educational research: cleaning up a messy construct. *Review of Educational Research, 307*-332.

Pareja, F. (1986). *La educación superior en el Ecuador* [The higher education in Ecuador]. Caracas: Regional Center for Higher Education in Latin America And the Caribbean (CRESALC)-UNESCO.

Pratt, D. D. (1992). Conceptions of teaching. *Adult Education Quarterly, 42*(4), 203-220.

Ribeiro, D. (1971). *La Universidad Latinoamericana [The Latin-American University]*. Caracas: Ediciones de la Biblioteca de la Universidad Central de Venezuela.

Richardson, J.T.E. (2005). Instruments for obtaining student feedback: a review of the literature. *Assessment & Evaluation in Higher Education*, *30*:4, 387-415. http://dx.doi.org/10.1080/02602930500099193

Richardson, J.T.E. (2012). The role of response bias in the relationship between students' perceptions of their courses and their approaches to studying in higher education. *British Educational Research Journal*, *38*, 399-418. https://doi.org/10.1080/01411926.2010.548857

Salvia, J., Ysseldyke, J. E., & Bolt, S. (2013). *Assessment in special and inclusive education* (12th ed.). Belmont, CA: Wadsworth.

Saroyan, A., & Amundsen, C. (2001). Evaluating university teaching: Time to take stock. *Assessment & Evaluation in Higher Education, 26*(4), 341-353. https://doi.org/10.1080/02602930120063493

Scherer R, Nilsen T., & Jansen M. (2016). Evaluating individual students' perceptions of instructional quality: An investigation of their factor structure, measurement invariance, and relations to educational outcomes. *Frontiers in Psychology*, *7*:110. https://doi.org/10.3389/fpsyg.2016.00110

Schmitt, T. A., Sass, D. A., Chappelle, W., & Thompson, W. (2018). Selecting the "best" factor structure and moving measurement validation forward: An illustration. *Journal of Personality Assessment*, *100*(4), 345-362. https://doi.org/10.1080/00223891.2018.1449116

Schön, D. A. (1983). *The reflective practitioner: how professionals think in action*. San Francisco: Jossey-Bass.

Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the Validity of Student Evaluation of Teaching: The State of the Art. *Review of Educational Research XX(X), pp. 1–45***.** http://dx.doi.org/10.3102/0034654313496870

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics.* Boston: Pearson.

Toland, M., & De Ayala, R.J. (2005). A multilevel factor analysis of students' evaluations of teaching. *Educational and Psychological Measurement*, *65*, 272-296. https://doi.org/10.1177/001316440426866

Uttl, B., White, C.A., & Gonzalez, D.W. (2017). Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related. *Studies in Educational Evaluation*, *54*, 22-42. https://doi.org/10.1016/j.stueduc.2016.08.007

Van der Linden, W. J. (2017). *Handbook of Item Response Theory, Volume Three: Applications.* New York, NY: Chapman and Hall/CRC.

Vandenberg, R. J., & Lance, C. E. (2000). A Review and synthesis of the measurement invariance literature: Suggestions , practices , and recommendations for organizational research. *Organizational Research Methods*, *3*(1), 4-70. https://doi.org/10.1177/109442810031002

Veas, A., Castejón, J. L., O'Reilly, C., & Ziegler, A. (2018). Mediation Analysis of the Relationship between Educational Capital, Learning Capital, and Underachievement among gifted secondary school students. Journal for the Education o the Gifted, 41(4), 369-385. https://doi.org/10.1177/0162353218799436

Wright, B. D. (1996). Local dependency, correlations and principal components. *Rasch Measurement Transactions, 10*(3), 509-511.

Wu, : L., Adams, R. J., & Wilson, M. R. (1998). ACER Conquest: Generalised item response modelling software [Computer software and manual]

Wu, M. L., Adams, R. J., Wilson, M. R., & Haldane, S. (2007). *ACER Conquest version 2.0: Generalised item response modelling software*. Camberwell, Victoria: Australian Council for Educational Research.

Xiao, Y., Liu, H., & Hau, K.-T. (2019). A comparison of CFA, ESEM, and BSEM in test structure analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, In press. https://doi.org/10.1080/10705511.2018.1562928

Young, K., Joines, J., Standish, T., & Gallagher, V. (2019) Student evaluations of teaching: the impact of faculty procedures on response rates, *Assessment & Evaluation in Higher Education*, *44* (1), 37-49, https://doi.org/10.1080/02602938.2018.1467878

Zabaleta, F. (2007). The use and misuse of student evaluation of teaching. *Teaching in Higher Education*, *12*, 55–76. https://doi.org/10.1080/13562510601102131

Ziegler, A., & Baker, J. (2013). Talent development as adaption: The role of educational and learning capital. In S. Phillipson, H. Stoeger, & A. Ziegler (Eds), *Exceptionality in East-Asia: Explorations in the actiotope model of giftedness* (pp. 18-39). Routledge.

Ziegler, A., Chandler, K., Vialle, W., & Stoeger, H. (2017). Exogenous and endogenous learning resources in the Actiotope Model of Giftedness and its significance for gifted education. *Journal for the Education of the Gifted, 40*, 310-333. https://doi.org/10.1177/0162353217734376

Ziegler, A., Debatin, T., & Stoeger, H. (2019). Learning resources and talent development from a systemic point of view. *Annals of the New York Academy of Sciences*, 1445, 39—51. https://doi.org/10.1111/nyas.14018

Ziegler, A., & Stoeger, H. (2017). Systemic gifted education. A theoretical introduction. *Gifted Child Quarterly, 61*, 183–193. doi:10.1177/0016986217705 713

Ziegler, A., Stoeger, H., & Balestrini, D. (2017). Systemic gifted education. In J. Riedl Cross, C. O'Rellly & T. Cross (Eds.), *Providing for the special needs of students with gifts and talents* (pp. 15-55). Dublin, Ireland: Kazoo Independent Publishing.

Corresponding author:
Alejandro Veas, PhD
University of Alicante.
Department of Developmental Psychology and Didactics
PO 03690
San Vicente del Raspeig, Alicante, Spain
E-mail: Alejandro.veas@ua.es

The effect of personal values on academic achievement

Juan F. Luesia[a], Milagrosa Sánchez-Martín[a], Isabel Benítez[b]

[a] Universidad Loyola Andalucía (Seville, Spain)
[b] University of Granada (Granada, Spain)

Abstract:
Numerous studies have demonstrated the importance of individual and contextual variables in explaining academic performance. Among the individual variables, personal characteristics such as sociodemographic variables have been investigated. Regarding the contextual variables, the influences of parental and school styles have been studied. However, personal values could also contribute to understanding students' achievement. The present study aims to analyse the relationship between personal values and academic performance. To do so, we first adapted two scales that aimed to measure the meaning of life and intellectual humility by following a committee approach. The Spanish version was administered to 54 students to evaluate the psychometric properties of the scales. The results and information provided by the experts were used to create revised versions of the scales, which were administered to 154 students together with other instruments focused on evaluating academic performance. The correlations between personal values and academic performance were computed, and the academic performance of students with different scores in personal values was compared. Intellectual humility was related to cognitive skills, and differences were identified in the academic performance between participants with high and medium scores in personal values. The implications and the utility of the adapted versions of the instruments are discussed.

Keywords:
personal values; academic performance; meaning of life; intellectual humility.

Introduction

Importance of personal values in University

The improvement of the educational system to help students to acquire not only knowledge but also global competencies for their personal and professional lives has been a main goal worldwide in recent decades (Reimers, 2009). However, academic performance has been mainly assessed by using only cognitive tests, which could be limited in terms of the representation of student capacities and tendencies. Therefore, in recent decades, researchers have focused their efforts on studying academic excellence by developing more comprehensive academic performance prediction models, where the evaluation of non-cognitive variables is considered.

In terms of research on students' achievements, the analysis of the variables affecting students' academic performance has been one of the key elements for proposing and developing educational strategies and educational policies. In fact, diverse approaches have achieved positive results related to the improvement of students' academic performance. For instance, the educational system in Finland focused on promoting teachers' professionalism and a school climate based on equity, flexibility, and creativity. The results indicate that the impacts of these two variables on students' performance was higher than those of other countries implementing strategies based on demanding learning standards (Sahlberg, 2007). Other researchers and practitioners have focused on investigating the relationship between schools' resources and students' performance, although according to the systematic review and meta-analyses conducted by Hanushek (1997), the direction and the strength of the connection between the two variables is not clear.

In attempts to deeply explore the variables influencing academic achievement, many individual characteristics have been studied. For instance, previous studies have confirmed the influence of sociodemographic characteristics on students' achievement. Thiele et al. (2016) found better academic results on students from wealthy areas and from the mainstream group. Muijs (1997) pointed to parental socioeconomic status (SES) and academic self-concept as good predictors of academic achievement. Personality traits seem to also be related to achievement. In a study by Diseth (2003), positive correlations occurred between both neuroticism and openness and achievement, but negative correlations were found between performance and agreeableness. These are only a few examples of studies addressing the analysis of the predictors of academic performance among the extensive literature on the topic. However, the influences of personal values on students' achievements has not yet been extensively studied.

The concept of personal values may include multiple variables. Globally, personal values are defined as a relatively stable belief in a particular mode of behaviour or consciousness, which is personally and socially preferable to other modes of behaviour (Rokeach, 1968). Individual values are grouped, forming a value system, defined as a permanent and stable organization over time that serves as a criterion to resolve conflicts and guide decision-making when more than one value is involved (Feather, 1972). Values are grouped into systems at the individual and social levels, in which one value determines and is deter-

mined by another. This implies a special interest in how much human values are found written in individuals' behaviour (Schwartz, 2006).

Previous studies have shown the importance of personal values in university learning and training. For instance, Matthews et al. (2007) indicated that values such as self-direction, self-aggrandizement or benevolent change may be related to university students' approaches to learning. However, the influence in terms of educational results that certain values may result in is unknown (Chase, et al., 2013). In other words, personal values could represent ideas or beliefs that go beyond specific situations since they represent indicators or criteria for evaluating the behaviours accepted by the society around us, although the direct impacts on achievement are not confirmed.

Theoretical models about personal values

Several theoretical models have considered personal values as an explanatory element of giftedness. Within the different approaches to giftedness, it has been suggested that considering a person to be talented should be associated with a superiority in socially desired personal values; this has been suggested, for example, by the pentagonal implicit theory of giftedness, in which the value criterion is a fundamental pillar in the development of giftedness (Sternberg & Zhang, 1995). Albert and Runco's model also specified that intelligence and performance are insufficient elements to define talent. The influences of family expectations, values and attitudes have motivating effects on talent (Albert & Runco, 1986). The model emphasizes the importance of the context where the subject operates as a determinant, along with other personal factors, of giftedness, which is defined by numerous factors beyond intelligence and performance. Similarly, Csikszentmihalyi and Robinson's approach to the conception of giftedness also details the influence that social context and culture can have on the development of giftedness (Csikszentmihalyi & Robinson, 2015).

Nevertheless, one of the most comprehensive models of giftedness is the actiotope model of giftedness, which focuses not only on personal attributes but also on the development of actions within a complex system (Ziegler, 2005; Ziegler & Stoeger, 2017; Ziegler et al., 2017). In this model, giftedness is considered an output of particularly effective actions. The model emphasizes the dynamic interaction of individuals with the environment. These actions are the consequence of three adaptations: biological adaptation, social adaptation and individual adaptation (Ziegler et al., 2013).

According to these models, personal values could play a role in understanding students' academic performance. In the actiotope model, they are considered an essential learning resource (Ziegler et al., 2019; Ziegler & Baker, 2013; Ziegler, Chandler et al., 2017). Among the personal values, the meaning of life and intellectual humility have been both defined and measured but not directly related to students' performance. The meaning of life refers to the belief that everyone has about the significance and the transcendence of their own life, which determine how people organize their experiences and how they plan the use of their energy (Steger & Frazier, 2005). Intellectual humility is a specific dimension of humility focused on the personal perception of intellectual strengths and

limitations and how they are handled (Davis et al., 2016). Both variables imply an individual evaluation of oneself and the development of strategies for pursuing personal goals, which might be related to academic achievement.

The present study aims to understand how personal values, specifically, the meaning of life and intellectual humility, are related to participants' characteristics and academic performance. To do so, we first adapted two Spanish scales originally created in English to measure the intended constructs (Study 1). Both scales were administered, and the psychometric properties were assessed. The results were used together with judgements from experts to implement changes in the scales, which led to a revised version. Then, a different group of participants responded to the two revised scales and some additional instruments focused on measuring cognitive and non-cognitive competencies (Study 2).

Method

Study 1: Adaptation and analysis of psychometric properties

Participants

A total of 54 participants in the last year of their bachelor's degrees (50 % women and 50 % men) responded to various instruments measuring personal values. The participants were asked to participate after completing a university pre-entrance assessment. Participation was voluntary, and they did not receive any compensation for their participation. Participants who agreed to participate in the study received a link to a web survey in which they took part after reading the information about the study and formally providing informed consent. Ethical approval for the research was obtained from the Universidad Loyola Andalucía Research Ethics Committee prior to the study.

Instruments

An assessment protocol was created to evaluate some personal values. The protocol included the following instruments:

The Spanish version of the Meaning of Life Questionnaire (MLQ-S)
The MLQ is a self-reported scale composed of ten items on the meaning of life that were answered using a seven-point Likert scale ranging from 1 (absolutely untrue) to 7 (absolutely true). The items measure two dimensions, the presence of meaning and search for meaning, and each dimension was composed of five items. The validation study of the original version reported adequate reliability of both subscales (Cronbach's alpha from .81 to .86 for presence of meaning and from .84 to .92 for search for meaning) and provided evidence of the validity supporting the internal structure of the scale and the expected relationships with other variables assessing well-being (Steger, Frazier, Oishi & Kaler, 2006). Details of the adaptation process and psychometric properties of the Spanish version administered to participants in Study 1 are described below.

The Spanish version of the Comprehensive Intellectual Humility Scale (CIHS-S)
The CIHS is a self-reported scale composed of 22 items evaluating four dimensions measuring intellectual humility: intellect and ego, openness to revising one's viewpoint,

respect for others' viewpoints, and lack of intellectual overconfidence. The items are statements that were assessed on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The validation studies conducted by the authors of the original scale did confirm the test-retest reliability of the scale and provided evidence of the validity supporting the utility of the scale for measuring the intended construct in both cases with multiple samples (Krumrei-Mancuso & Rouse, 2016). Details of the adaptation process and psychometric properties of the Spanish version administered to participants in Study 1 are described below.

Additional items measuring self-criticism

A pool of additional items was included as part of the evaluation with the aim of covering additional dimensions that should be theoretically part of the intended constructs in Spanish students. Specifically, self-criticism, although partly included in the CIHS as part of the dimension on openness to revisiting one's viewpoint, was only measured as the ability to self-critic when receiving external output. However, self-knowledge about oneself and others and the ability to evaluate oneself were also considered relevant, especially for the academic context. For that reason, some additional items were included as part of the assessment protocol. First, we adapted three items from the Social Responsibility Scale (SRS, Ramos, Armentia, & de la Fuente, 2008). The SRS is a scale created to evaluate the changes in the social responsibility of students attending a course on the topic. Three items were selected specifically because they measured self-criticism and were adapted to capture general behaviours (instead of changes related to attending a training course as they were designed). Second, we created four ad hoc items focused on self-criticism indicators not covered by previous instruments.

Procedure

The MLQ and CIHS were adapted by following a committee approach, as Harkness and Schoua-Glusberg (1998) described. First, three independent translators generated a version of each scale. Then, the translation coordinator compared the three versions and identified any discrepancies. The discrepancies were discussed in a consensus meeting where the final version was agreed upon. To unify the instrument format and use the most common response scale, the committee members suggested using the same response scale for all the items. Therefore, a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree) was consistently adopted. Appendixes 1 and 2 include the original and adapted versions of the items in the MLQ and CIHS (MLQ-S and CIHS-S), respectively. Additional items measuring self-criticism were reviewed by two members of the research team. All the additional items included in the assessment protocol are listed in Appendix 3 (self-criticism items). Both the scales and the additional items were included in a web survey that was created using Qualtrics (https://www.qualtrics.com). The link to the survey was provided to participants who responded during a virtual session. After giving the instructions, the participants completed the task.

Concurrently, three experts on personal values were contacted, and they were asked to evaluate the adapted versions of the scales and items measuring self-criticism. The experts' task consisted of three phases:

1) evaluation of the theoretical definitions of the constructs, 2) evaluation of the adequacy of the items, and 3) formulation of suggestions for improving the items and scales. The information provided by experts was used to interpret the psychometric results obtained from participants' responses and to propose modifications to the instruments.

Analysis

The responses of participants were used to analyse the psychometric properties of the scale. First, reliability was assessed by computing Cronbach's alpha. Item-test correlations were computed to explore the properties of each item. Exploratory factor analysis (EFA) was conducted to evaluate the dimensionality of the instruments. Principal-axis factor analysis with oblique rotation (promax) was used to validate the original instruments. Data analysis was conducted using SPSS Statistics (version 26).

Study 2: Values and academic competencies

Participants

A total of 154 first-year university students (35 % women; 65 % men) responded to a booklet including the revised versions of the instruments from Study 1 and additional tests and questions measuring academic performance. The average age of the sample was 17.76 years (SD = 1.16). Participants were asked to voluntarily participate at the university. Participants who agreed to participate in the study received a link to a web survey in which they took part after reading the information about the study and formally providing informed consent. Ethical approval for the research was obtained from the Universidad Loyola Andalucía Research Ethics Committee prior to the study.

Instruments

Personal Values

The participants responded to the revised versions of the instruments obtained from Study 1: MLQ-S-R and CIHS-S-R.

Ability Battery BAT-7 (Arribas-Águila et al., 2013).

The BAT-7 is one of the most commonly used instruments for assessing cognitive abilities. The battery measures seven competencies: verbal ability (V), spatial ability (E), attention (A), reasoning (R), mathematical ability (N), mechanical aptitude (M) and spelling (O). These competencies are measured by seven subtests that have previously been shown to have adequate psychometric properties (Cronbach's alphas between .78 and .95; Sánchez-Sánchez & Arribas-Águila, 2014). According to the sample age (16 to 18 years old), the higher level version of the instrument (S) was used.

Previous performance: Grades achieved in high school (HSGPA) and grades obtained on the national University Entrance Examinations (UEE) were also collected.

Procedure and analysis

Data were collected through the web survey Qualtrics (https://www.qualtrics.com). The link to the survey was provided to participants who responded during a virtual session with videoconferencing through the Webex platform during which researchers were available for questions or technical

problems. SPSS Statistics (version 26) was used for the analyses. First, bivariate correlations between academic achievement scores and personal values were analysed. Then, we explored differences between participants by grouping them according to their scores on the MLQ-S-R and CIHS-S-R. To achieve this, a classification based on terciles was established by using the total and dimension scores on both scales. Participants were divided into three profiles: low, medium and high. Considering these groups, a bivariate ANOVA was conducted using Snedecor's F statistic and the Honestly Significant Difference (HSD), as recommended by Field (2009). A post hoc test was performed to compare the mean differences in the academic performance between participants with different profiles based on personal values.

Results

Study 1: Psychometric properties of the adapted versions

Psychometric properties

Table 1 shows the psychometric properties of the MLQ-S and CIHS-S and the additional items used to measure self-criticism.

As Table 1 indicates, the MLQ showed adequate properties. The reliability of the presence of meaning subscale was .84, and item-test correlations reached values ranging from .56 to .80. In addition, the alpha did not increase when removing any of the items. The search for meaning subscale obtained a Cronbach's alpha value of .84. The items' properties also indicated adequate values, although removing item 2 would increase the reliability of the scale.

Regarding the CIHS, three dimensions reached adequate reliability indexes, but the lack of intellectual overconfidence dimension obtained a value lower than .7 (α = .66). Item 5 was the item with the poorest properties as the item-test correlation was medium (r = .26), and the alpha increased when removing it. The other three dimensions and the items composing them achieved adequate values. The internal consistency of the complete scale did not reach adequate values.

The self-criticism items worked adequately as a scale. The Cronbach's alpha confirmed the internal consistency of the scale (α = .71), and all the items exhibited good properties. Item 3 reached the lowest correlation with the total score, and removing it improved the stability of the scale.

Dimensionality

Table 2 shows the factor pattern when conducting the principal-axis analysis with promax rotation for the MLQ. The Kaiser-Meyer Olkin measure of sampling adequacy suggested that the sample was factorable (KMO = .758), and Bartlett's test of sphericity was significant (χ^2_{45} = 247.332, p = .000).

As Table 2 shows, the factor pattern of the MLQ fit the theoretical subscales. The solution with the two factors reaching eigenvalues higher than 1 explained a total of 63.31 % of the variance. The items' loadings reflected the theoretical distribution proposed for the original version of the scale.

The results for the CIHS showed that the sample was factorable (KMO = .703), and Bartlett's test of sphericity was significant (χ^2_{231} = 536.584, p = .000). Table 3 shows the factor pattern when conducting the principal-axis analysis with promax rotation.

Table 1 Psychometric properties of the MLQ-S and CIHS-S

Scale	Subscale	Items	Discrimination Index (Total Item-test correlation)	Alpha if removed
Meaning of Life	Presence of Meaning (α=.842)	1	.555	.834
		4	.641	.812
		5	.621	.818
		6	.799	.772
		9*	.649	.812
	Search for Meaning (α=.836)	2	.492	.839
		3	.626	.806
		7	.714	.780
		8	.663	.795
		10	.695	.786
Intellectual humility (α=.786)	Lack of Intellectual Overconfidence (*α=.664*)	1*	.445	.608
		2*	.560	.586
		3*	.347	.639
		4*	.499	.582
		5*	.257	.682
		12*	.359	.636
	Openness to Revising One's Viewpoint (α=.883)	6	.685	.870
		7	.766	.847
		8	.761	.848
		9	.744	.857
		10	.676	.868
	Respect for Others' Viewpoints (α=.861)	11	.713	.826
		13	.599	.848
		14	.699	.829
		15	.565	.853
		19	.673	.834
		20	.679	.834
	Independence of Intellect and Ego (α=.788)	16*	.523	.762
		17*	.650	.721
		18*	.704	.697
		21*	.524	.761
		22*	.433	.787
Self-criticism (α=.710)		1	.500	.658
		2	.368	.689
		3	.247	.724
		4	.278	.707
		5	.564	.642
		6	.569	.641
		7	.468	.666

*Reversed items.

Table 2 Factor pattern of principal-axis factor analysis of items of the MLQ-S

Scale	Subscale	Items	Item loading
Meaning of Life	Factor 1 (40.56 % of explained variance)	1	.633
		4	.682
		5	.680
		6	.869
		9*	.688
	Factor 2 (22.75 % of explained variance)	2	.534
		3	.664
		7	.809
		8	.827
		10	.759

*Reversed items

Table 3 Factor pattern of principal-axis factor analysis of items of the CIHS-S

Item	Factor						
	1	2	3	4	5	6	7
1				.970			
2				.748			
3						.556	
4						.914	.393
5							.482
6	.683						
7	.800						
8	.966						
9	.707						
10	.680						
11		.635					
12		.315				.327	.450
13		.658					.339
14		.810					
15		.542					
16			.533				
17			.911				
18			.707				
19		.699					
20		.742					
21					.743		
22					.862		

As Table 3 indicates, the original solution proposed seven factors explaining 75.34 % of the variance. The distribution of the items in the factors was similar to the theoretical structure proposed for the original version. Factor 1 explaining 27.1 % of the variance included items in the openness to revising one's viewpoint. Factor 3 explained 12.13 % of the variance and was represented by items measuring respect for others' viewpoints. However, items in both dimensions, the lack of intellectual overconfidence and the independence of intellect and ego, were divided into different factors: factors 3 and 5 for the former and factors 4, 6 and 7 for the latter.

The results for the self-criticism scale showed that the sample was not factorable (KMO = .698), but the sphericity requirement was reached as Bartlett's test of sphericity was significant (χ^2_{21} = 71.196, p = .000). Table 4 shows the factor pattern when conducting the principal-axis analysis with promax rotation.

As Table 4 indicates, the original solution proposed two factors explaining 53.69 % of the variance. The distribution of the items in the factors did not respond to the theoretical criteria, and items' content did not seem to be grouped according to what they were measuring.

Modified version

Previous results indicated problematic elements in the adapted versions of the scales and in the items used for assessing self-criticism. Therefore, experts' judgements were reviewed and used to propose modifications to the items. Table 5 summarizes the main contributions of the experts.

As Table 5 shows, the experts indicated that the content of the items in the MLQ-S was redundant and suggested reducing it. Specifically, the experts proposed discarding three items. Although these items showed adequate psychometric properties, they were removed in order to obtain a more parsimonious version of the scale.

Experts also raised suggestions for the CIHS-S. First, they proposed reformulating some items in the lack of intellectual overconfidence subscale because of the complexity of their terms and expressions. Due to the low reliability found for that subscale, some modifications were implemented to simplify the items' structures. In addition, items 8 and 10 in the openness to revisiting one's viewpoint subscale were viewed as repetitive. The experts suggested removing these items and merging that dimension with the self-criticism scale. As both dimensions were intended to measure the same construct and self-criticism could be part

Table 4 Factor pattern of principal-axis factor analysis of items of the self-criticism scale

Scale	Subscale	Items	Item loading
Meaning of Life	Factor 1 (38.51 % of explained variance)	1	.639
		4	.335
		6	.888
		7	.511
	Factor 2 (15.17 % of explained variance)	2	.708
		3	.301
		5	.616

Table 5 Experts' suggestions about items and scales

Scale	Subscale	Items	Comments	Suggestion
Meaning of life	Search for Meaning	2	The content is ambiguous, and the indicator is already measured in item 7.	Remove the item
		3	The content is ambiguous. The term "always" is extreme, and the indicator is already measured in item 8.	
		10	The indicator is already measured in other items of the scale and the past continuous could be confusing.	Modify the item
Intellectual humility	Lack of Intellectual Overconfidence	3*	The item is difficult to understand because of the expressions "not very likely" and "incorrect idea". The formulation includes negations.	Modify the item
		5*	The item is difficult to understand because of the expressions "not very likely" and "influence my ideas". The formulation includes negations.	Modify the item
		12*	The item includes confusing expressions such as "rarely".	Modify the item
	Openness to Revising One's Viewpoint	8	The item repeats a concept evaluated in item 7.	Remove the item
		10	The item repeats a concept evaluated in item 10.	Remove the item

of the concept named intellectual humility, we created a new dimension composed of the three original items in the CIHS (items 6, 7 and 9) and the six items included to assess self-criticism. Although item 3 in the self-criticism scale did not achieve good psychometric properties, no suggestions were collected to modify it. Therefore, we retained the previous version as part of the final test. Appendixes 4 and 5 include the items included in the revised versions of the MLQ-S and CIHS-S, named MLQ-S-R and CIHS-S-R, respectively.

Study 2: Values and academic competencies

Psychometric properties

Both revised versions achieved adequate psychometric properties. The two dimensions of the MLQ-S-R obtained Cronbach's alphas higher than .7. Specifically, the presence of meaning dimension reached a value of .81, and the search for meaning dimension composed of two items of the original version (items 7 and 8 of MLQ-S) and one item with modifications (item 10 of MLQ-S) obtained a value of .71. Item-test correlation ranged from .44 to .69 in the first dimension and from .48 to .60 in the second dimension.

In the CIHS-S-R, the psychometric properties reflected an improvement in the internal consistency. The global Cronbach's alpha was .93, with four subscales obtaining

higher values than in the previous version. Changes in the lack of intellectual overconfidence subscale made those items work better, exhibiting correlations with total scores ranging from .49 to .66 and a Cronbach's alpha of the subscale of .82. The two dimensions without modifications also increased previous alphas with current values of .88 and .81 for respect for others' viewpoints and independence of intellect and ego, respectively. The last dimension merging items from the previous openness to revising one's viewpoint and self-criticism items reached a Cronbach's alpha of .89. Item-test correlations ranged from .36 to .73. To better reflect the content of the dimension, we renamed it the dimension openness to revising one's self-knowledge.

In terms of dimensionality, the MLQ-S-R achieved the expected factors, although item 10 received high loadings in both factors (.59 in the presence of meaning and .56 in the search for meaning). The items of the CIHS-S-R did fit the theoretical structure of the original in terms of items achieving the highest loading on the expected factor, although the factor loadings of most the items were high for all the factors pointing to a unidimensional structure explaining 38.86 % of the variance or a two-dimensional solution accounting for 50.88 % of the variance. A two-factor solution would split items into two parts: A) items from the lack of intellectual overconfidence and independence of intellect and ego dimensions, and B) items from respect for others' viewpoints and openness to revising self-knowledge. Examining the content of the items, part A would include items focused on the strength of one's own arguments, and part B would be more related to the ability to value and consider external opinions and to self-knowledge. In order to retain the theoretical subscales of the CIHS, we considered the original structure with four dimensions treated as subscales for the following analysis.

Personal values and academic performance

The correlations between personal values and cognitive skills are shown in Table 6.

As Table 6 indicates, negative correlations were found between the CIHS-S-R and BAT-7 total scores ($r = -.19$, $p = .03$) and between the CIHS-S-R total score and the cognitive dimension of attention ($r = -.19$, $p = .03$). Considering the four dimensions of the CIHS-S-R, only the openness to revising self-knowledge dimension achieved significant results ($r = -.19$, $p = .03$). Verbal ability also had a negative and significant correlation with that dimension ($r = -.17$, $p = .05$). No other significant correlations were found between personal values and academic performance, cognitive abilities or previous achievement obtained in high school and on the UEE, although the correlations between dimensions of personal values were as expected.

To deeply explore the relationships between personal values and academic performance, students' profiles were investigated. Participants were divided into three groups in each dimension of the MLQ-R-S and CIHS-S-R scales. Tables 7 and 8 show the results from the ANOVA conducted to compare the means of the participants with low, medium and high scores on each of the dimensions and subscales.

Table 6 Correlations between personal values and academic performance

	Variable	*n*	*M*	*SD*	[1]	[2]	[3]	[4]	[5]	[6]	[7]
1.	MLQ-S-R1 [1]	148	17.14	3.86	1						
2.	MLQ-S-R2 [2]	148	9.97	2.65	.42**	1					
3.	CIHS-S-R (Total score) [3]	147	-3.59	7.23	.24**	.35**	1				
4.	CIHS-S-R (Lack) [4]	148	2.32	4.50	.12	-.01	-.55**	1			
5.	CIHS-S-R (Respect) [5]	147	11.24	2.21	.41**	.36**	.47**	.37**	1		
6.	CIHS-S-R (Ego) [6]	148	22.85	4.50	.35**	.31**	.45**	.38**	.55**	1	
7.	CIHS-S-R (Openness) [7]	148	17.34	4.13	.15	-.00	-.37**	.64**	.87**	.50**	1
8.	Verbal ability	137	21.82	4.10	-.08	.05	-.12	-.06	-.11	.02	-.17*
9.	Spatial ability	137	18.54	5.28	.00	.129	-.14	.09	-.04	.03	-.09
10.	Attention	137	38.01	1.71	-.01	.00	-.19*	.09	-.09	.07	-.09
11.	Reasoning	137	19.79	4.33	-.01	.02	-.10	.09	-.01	.02	-.08
12.	Mathematical ability	137	16.95	5.01	-.08	.11	-.07	.00	-.09	-.08	-.14
13.	Mechanical aptitude	137	18.42	3.74	-.07	.06	-.07	.05	-.02	.00	-.06
14.	Spelling	137	22.05	5.09	-.05	.10	-.10	.010	-.03	.02	-.143
15.	BAT-7 Total Score	137	356.52	81.05	-.08	.12	-.19*	.054	-.10	.01	-.19*
16.	HSGPA	128	7.95	1.04	-.01	.08	-.10	-.01	.00	.14	-.08
17.	UEE	128	7.67	1.07	.00	.13	-.07	.05	.05	.12	.00

*p<.05. **p<.01

MLQ-S-R: Meaning of Life Questionnaire. MLQ-S-R1: Presence of Meaning of life. MLQ-S-R2: Search of Meaning of life. CIHS-S-R: Intellectual Humility. HSGPA: Grades achieved in High School. UEE: University Entrance Examinations.

Table 7 Means differences on cognitive skills and previous achievement between MLQ-R-S profiles

	MLQ-S-R:Presence of meaning	MLQ-S-R:Search for meaning
Verbal ability	$F(2,134) = 2{,}27$ $p = .11$	$F(2,134) = 0.56$ $p = .57$
Spatial ability	$F(2,134) = 0.16$ $p = .85$	$F(2,134) = 1.12$ $p = .33$
Attention	$F(2,134) = 0.07$ $p = .93$	$F(2,134) = 0.32$ $p = .73$
Reasoning	$F(2,134) = 0.51$ $p = .60$	$F(2,134) = 0.09$ $p = .91$
Mathematical ability	$F(2,134) = 1.44$ $p = .24$	$F(2,134) = 2.71$ $p = .07$
Mechanical aptitude	$F(2,134) = 0.32$ $p = .73$	$F(2,134) = 0.88$ $p = .42$
Spelling	$F(2,134) = 0.82$ $p = .44$	$F(2,134) = 0.93$ $p = .40$
BAT-7 Total Score	$F(2,134) = 1.08$ $p = .34$	$F(2,134) = 1.82$ $p = .17$
HSGPA	$F(2,134) = 1.58$ $p = .21$	$F(2,134) = 0.97$ $p = .38$
UEE	$F(2,134) = 0.90$ $p = .41$	$F(2,134) = 0.82$ $p = .44$

MLQ-S-R: Meaning of Life Questionnaire. MLQ-S-R1: Presence of Meaning of life. MLQ-S-R2: Search of Meaning of life. CIHS-S-R: Intellectual Humility. HSGPA: Grades achieved in High School. UEE: University Entrance Examinations.

Table 8 Means differences on cognitive skills and previous achievement between CIHS-R-S profiles

	CIHS-S-R: Total Score	**CIHS-S-R: Lack**	**CIHS-S-R: Respect**	**CIHS-S-R: Ego**	**CIHS-S-R: Openness**
Verbal ability	*F(2,132)* = 0.80 *p* = .45	*F(2,134)* = 2.29 *p* = .11	*F(2,134)* = 1.01 *p* = .37	*F(2,132)* = 1.36 *p* = .26	*F(2,132)* = 2.20 *p* = .11
Spatial ability	*F(2,132)* = 2.09 *p* = .13	*F(2,134)* = 0.24 *p* = .79	*F(2,134)* = 1.04 *p* = .36	*F(2,132)* = 0.52 *p* = .59	*F(2,132)* = 0.44 *p* = .64
Attention	*F(2,132)* = 5.15 *p* = .01$^{H<L;\ H<M}$	*F(2,134)* = 0.14 *p* = .87	*F(2,134)* = 1.14 *p* = .32	*F(2,132)* = 0.10 *p* = .91	*F(2,132)* = 1.39 *p* = .25
Reasoning	*F(2,132)* = 2.85 *p* = .06$^{H<M}$	*F(2,134)* = 0.70 *p* = .50	*F(2,134)* = 0.52 *p* = .60	*F(2,132)* = 0.43 *p* = .65	*F(2,132)* = 0.44 *p* = .65
Mathematical ability	*F(2,132)* = 0.11 *p* = .90	*F(2,134)* = 1.06 *p* = .35	*F(2,134)* = 0.20 *p* = .82	*F(2,132)* = 2.74 *p* = .07	*F(2,132)* = 1.60 *p* = .21
Mechanical aptitude	*F(2,132)* = 0.69 *p* = .50	*F(2,134)* = 0.01 *p* = .99	*F(2,134)* = 0.04 *p* = .96	*F(2,132)* = 0.27 *p* = .76	*F(2,132)* = 0.62 *p* = .54
Spelling	*F(2,132)* = 3.19 *p* = .04$^{H<M}$	*F(2,134)* = 0.92 *p* = .40	*F(2,134)* = 1.36 *p* = .26	*F(2,132)* = 2.28 *p* = .11	*F(2,132)* = 1.06 *p* = .35
BAT-7 Total Score	*F(2,132)* = 3.93 *p* = .02$^{H<M}$	*F(2,134)* = 0.38 *p* = .68	*F(2,134)* = 0.89 *p* = .41	*F(2,132)* = 0.37 *p* = .69	*F(2,132)* = 2.43 *p* = .09$^{H<L}$
HSGPA	*F(2,132)* = 1.6 *p* = .21	*F(2,134)* = 3.22 *p* = .04$^{H<M}$	*F(2,134)* = 0.00 *p* = .99	*F(2,132)* = 0.14 *p* = .87	*F(2,132)* = 0.12 *p* = .89
UEE	*F(2,132)* = 0.96 *p* = .38	*F(2,134)* = 0.64 *p* = .53	*F(2,134)* = 0.03 *p* = .97	*F(2,132)* = 0.20 *p* = .82	*F(2,132)* = 0.09 *p* = .92

MLQ-S-R: Meaning of Life Questionnaire. MLQ-S-R1: Presence of Meaning of life. MLQ-S-R2: Search of Meaning of life. CIHS-S-R: Intellectual Humility. HSGPA: Grades achieved in High School. UEE: University Entrance Examinations.

As Table 7 indicates, no significant differences were found between the profiles of the MLQ-R-S dimensions and academic performance. Mathematical ability was the only cognitive skill where a trend was identified with the MLQ-S-R2-high group with a higher mean (M = 18.49, SD = 4.56) than the MLQ-S-R2-medium group (M = 16.07, SD = 4.95).

When comparing groups using CIHS-S-R total scores, significant differences were found in BAT-7 total scores and in both spelling and attention dimensions (see Table 8). With respect to the BAT-7 total scores, the mean obtained by the high group (M = 330.74, SD = 82.87) was significantly lower than that of the medium group (M = 377.55, SD = 76.71). These differences between these two groups were also significant in the spelling dimension (high group: M = 20.79, SD = 5.02; medium group: M = 23.55, SD = 5.34). Regarding the attention dimension, the mean score of the high group (M = 33.79, SD = 10.69) was significantly lower than that of the low group (M = 40.02, SD = 11.06) and medium group (M = 39.86, SD = 9.77). In addition, the results obtained for the reasoning dimension tended towards significance in the same direction as the above results (high group: M = 19.00, SD = 3.85; medium group: M = 21.05, SD = 4.14).

Multiple comparisons considering the different groups (high/medium/low) of the four CIHS-S-R subscales yielded different results. When comparing the CIHS-S-R-lack groups, there were significant differences

only in HSGPA. The results showed that the scores obtained by participants in the high group (M = 7.63, SD = 0.95) were significantly lower than those obtained by participants in the medium group (M = 8.21, SD = 1.05). Regarding the remaining factors of intellectual humility, no significant differences were detected between groups.

Discussion

The present study aimed to understand how personal values are related to participants' characteristics and academic performance. The final goal of the study was to evaluate whether personal values play a role in the definition of academic excellence.

According to the results, personal values assessed in the present study seem to not be related to academic performance. First, the correlations between personal values and the indicators of academic performance were mostly not significant. Only the total score of the CIHS-S-R reached significant correlations with the total score on the BAT-7, but these correlations were negative, which suggests that higher intellectual humility is related to lower cognitive skills. A similar pattern is found between the CIHS-S-R dimension of openness to revising one's viewpoint. One possible explanation for this lack of relationship may be due to the non-identification of specific thinking styles in students, such as cognitive style (Sagiv et al., 2013) or strategic thinking (Steptoe-Warren, 2011), which could provide further evidence on the association between personal values and academic performance. Second, the profile exploration also shows that connection as significant differences were found in the means of the cognitive skills between groups divided by their total scores on the CIHS-S-R. Differences in attention, reasoning, spelling, and total scores on the BAT-7 indicated worse performance among participants with higher levels of intellectual humility. However, medium levels of intellectual humility were connected to higher scores in these cognitive skills. In terms of the indicators of academic performance, the results showed that participants with higher HSGPAs were in the medium group of the lack of intellectual overconfidence dimension of the CIHS-S-R. In this sense, previous studies have shown that giftedness is not a unitary measure of intelligence (Sternberg & Zhang, 1995) so that as it increases, other talents become more relevant, such as creativity (Runco, 2005). These results suggest that intellectual humility might have an optimal score that is located in the middle of the range.

Furthermore, meaning of life was unrelated to academic performance, which suggests that this personal value does not have a direct relationship with academic performance. Therefore, contrary to expectations, the meaning of life and intellectual humility are not clearly related to academic performance in students with a mean age of 17.76 years, but this is something that could change over the years. This could be explained by the students' own life stage. The learning environment is in a process of change as students move from high school to university (Ziegler et al., 2013). This may imply that personal values have a greater influence when students are studying at university and may even have a direct impact on GPA (Harackiewicz et al., 2018).

Considering the results in Study 2, three conclusions are drawn. First, perhaps the question is not whether personal values are related to academic performance; otherwise, the objective must be oriented to

determine which personal values could be associated with academic performance. For instance, Matthews et al. (2007) suggested that personal values such as self-direction, self-aggrandizement or benevolent change may be related to learning. Second, personal values may not be directly connected to academic performance, but other noncognitive variables could moderate the relationship. For instance, previous studies show that these variables include personality traits, motivation factors, self-regulatory learning and other personal qualities (Duckworth et al. 2015; Richardson et al., 2012). Finally, the BAT-7 and indicators of academic performance may not capture the abilities of students to be successful in their professional lives. Although the use of batteries for assessing cognitive skills is extensive, the tests used may not capture the actual capacity of students to handle specific challenges at university. Previous studies have shown that broad assessments of student abilities can predict subsequent performance to a greater extent than if only purely cognitive or intellectual variables were considered (Niessen et al., 2018; Schmitt et al., 2009). Future studies will address a more extensive evaluation of students where additional instruments measuring personal values and other noncognitive variables are included as part of the assessment protocol. In addition, other indicators of academic performance such as achievement during university courses will be incorporated. Furthermore, future data collection should include participants with a wider range of ages in order to investigate whether personal values play a relevant role in different age groups.

Besides the substantive results, the present study also provides two reliable assessment instruments to measure the meaning of life and intellectual humility in Spanish participants. In this study, we adapted two instruments to Spanish and analysed the psychometric properties. In terms of reliability, both instruments showed adequate properties. In terms of dimensionality, the results suggest that intellectual humility is not defined in Spanish participants as it is theoretically. In fact, some additional indicators were incorporated as part of the construct since they were relevant in the specific population. Self-criticism and self-knowledge are elements that have acquired importance in Spanish university teaching-learning processes (Abad-Segura, 2019; Fidalgo & García, 2009). Future studies should be developed to confirm the structural dimensionality of the CIHS-S-R in order to determine whether subscales should be considered or whether total scores are better for reflecting the nature of the construct, as was considered in the present study. Self-criticism and self-knowledge are elements that have acquired importance in university teaching-learning processes.

Although our findings indicate that personal values are not directly related to academic performance, they are already relevant for educational purposes. Therefore, they should be evaluated properly by using adequate instruments, and new studies should shed light on the impact they have in people's lives.

References

Abad-Segura, E., & González-Zamar, M. D. (2019). Análisis de las competencias en la educación superior a través de flipped classroom. *Revista Iberoamericana De Educación, 80*(2), 29-45. https://doi.org/10.35362/rie8023407

Albert, R.S. & Runco, M.A. (1986). The activament of eminence: a model based on a longitudinal study of exceptionally gifted boys and their families. In R.J. Stemberg and J.E. Davidson (Eds.), *Conceptions of giftedness,* (pp. 332-357). Cambridge University Press.

Arribas-Águila, D., Santamaría, P., Sánchez-Sánchez, F., & Pinto, I. (2013). *BAT-7. Batería de Aptitudes de TEA, Manual.* Madrid, Spain: TEA Ediciones.

Chase, J. A., Houmanfar, R., Hayes, S. C., Ward, T. A., Vilardaga, J. P., & Follette, V. (2013). Values are not just goals: Online ACT-based values training adds to goal setting in improving undergraduate college student performance. *Journal of Contextual Behavioral Science, 2*(3–4), 79–84. https://doi.org/10.1016/j.jcbs.2013.08.002

Csikszentmihalyi, M., & Robinson, R. (2015). Culture, Time, and the Development of Talent. In M. Csikszentmihalyi (Ed.), *The systems model of creativity: The collected works of Mihaly Csikszentmihalyi* (pp. 27–46). Springer, Dordrecht. https://doi.org/10.1007/978-94-017-9085-7_3.

Davis, D. E., Rice, K., McElroy, S., DeBlaere, C., Choe, E., Van Tongeren, D. R., & Hook, J. N. (2016). Distinguishing intellectual humility and general humility. *The Journal of Positive Psychology*, *11*(3), 215-224. https://doi.org/10.1080/17439760.2015.1048818

Diseth, Å. (2003). Personality and approaches to learning as predictors of academic achievement. *European Journal of personality*, *17*(2), 143-155. https://doi.org/10.1002/per.469

Duckworth, A. L., & Yeager, D. S. (2015). Measurement Matters: Assessing Personal Qualities Other Than Cognitive Ability for Educational Purposes. *Educational Researcher, 44*(4), 237–251. https://doi.org/10.3102/0013189X15584327

Feather, N. T. (1972). Value Systems and Education: The Flinders Programme of Value Research. *Australian Journal of Education*, *16*(2), 136–149. https://doi.g/10.1177/000494417201600204

Fidalgo, R., & García, J.N. (2009). La evaluación de la metacognición en la composición escrita. *Estudios de Psicología, 30*(1), 51–72. https://doi.org/10.1174/021093909787536290

Field, A. P. (2009). *Discovering statistics using SPSS: (and sex and drugs and rock 'n' roll)*. London, Royaume-Uni: SAGE.

Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational evaluation and policy analysis*, *19*(2), 141-164. https://doi.org/10.3102/01623737019002141

Harackiewicz, J. M., & Priniski, S. J. (2018). Improving Student Outcomes in Higher Education: The Science of Targeted Intervention. *Annual Review of Psychology, 69*, 409–435. https://doi.org/10.1146/annurev-psych-122216-011725

Harkness, J. A., & Schoua-Glusberg, A. (1998). Questionnaires in translation. *ZUMA-Nachrichten Spezial, 7*, 87–125.

IBM Corp. (2019). *IBM SPSS Statistics for Windows* (Version 26.0). Armonk, NY: IBM Corp. [Computer software]

Krumrei-Mancuso, E. J., & Rouse, S. V. (2016). The development and validation of the comprehensive intellectual humility scale. *Journal of Personality Assessment*, *98*(2), 209-221. https://doi.org/10.1080/00223891.2015.1068174

Matthews, B., Lietz, P., & Darmawan, I. G. N. (2007). Values and learning approaches of students at an international university. *Social Psychology of Education, 10*(2), 247–275. https://doi.org/10.1007/s11218-007-9019-x

Muijs, R. D. (1997). Predictors of academic achievement and academic self-concept: A longitudinal perspective. *British Journal of Educational Psychology*, *67*(3), 263-277. https://doi.org/10.1111/j.2044-8279.1997.tb01243.x

Niessen, A. S. M., Meijer, R. R., & Tendeiro, J. N. (2018). Admission testing for higher education: A multi-cohort study on the validity of high-fidelity curriculum-sampling tests. *PLOS ONE, 13*(6), e0198746. https://doi.org/10.1371/journal.pone.0198746

Ramos, J. M. G., Armentia, P. G., & de la Fuente, M. O. (2008). Validación y medida de la responsabilidad social en la universidad. *Revista Complutense de Educación*, *19*(2), 385-404.

Reimers, F. (2009). Educating for global competency. In J. E. Cohen & M. B. Malin (Eds.), *International perspectives on the goals of universal basic and secondary education* (pp-183-202). New York: Routledge.

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: *A systematic review and meta-analysis. Psychological Bulletin, 138*(2), 353–387. https://doi.org/10.1037/a0026838

Rokeach, M. (1968). The Role of Values in Public Opinion Research. *The Public Opinion Quarterly*, *32*(4), 547–559. https://doi.org/10.1086/267645

Runco, M. A. (2005). Creative Giftedness. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (p. 295–311). Cambridge, England: Cambridge University Press. https://doi.org/10.1017/CBO9780511610455.017

Sagiv, L., Amit, A., Ein-Gar, D., & Arieli, S. (2014). Not all great minds think alike: Systematic and intuitive cognitive styles. *Journal of Personality, 82*(5), 402–417. https://doi.org/10.1111/jopy.12071

Sahlberg, P. (2007). Education policies for raising student learning: The Finnish approach. *Journal of education policy*, *22*(2), 147-171. https://doi.org/10.1080/02680930601158919

Sánchez-Sánchez, F., &Arribas-Águila D. (2014). BAT-7, Batería de Aptitudes de TEA: Descripción y datos psicométricos. *International Journal of Developmental and Educational Psychology, 2*(1), 353-364. https://doi.org/10.17060/ijodaep.2014.n1.v2.450

Schmitt, N., Keeney, J., Oswald, F. L., Pleskac, T. J., Billington, A. Q., Sinha, R., & Zorzie, M. (2009). Prediction of 4-Year College Student Performance Using Cognitive and Noncognitive Predictors and the Impact on Demographic Status of Admitted Students. *Journal of Applied Psychology, 94*(6), 1479–1497. https://doi.org/10.1037/a0016810

Schwartz, S. H. (2006). Basic human values: Theory, measurement, and applications. *Revue française de sociologie*, *47*(4), 929-968. https://doi.org/10.3917/rfs.474.0929

Steger, M. F., & Frazier, P. (2005). Meaning in life: One link in the chain from religiousness to well-being. *Journal of Counseling Psychology, 52,* 574–582. https://doi.org/10.1037/0022-0167.52.4.574

Steger, M. F., Frazier, P., Oishi, S., & Kaler, M. (2006). The meaning in life questionnaire: Assessing the presence of and search for meaning in life. *Journal of counseling psychology*, *53*(1), 80-93. https://doi.org/10.1037/0022-0167.53.1.80

Steptoe-Warren, G., Howat, D., & Hume, I. (2011). Strategic thinking and decision making: literature review. *Journal of Strategy and Management, 4*(3), 238–250. https://doi.org/10.1108/17554251111152261

Sternberg, R. J., & Zhang, L. (1995). What Do We Mean by Giftedness? A Pentagonal Implicit Theory. *Gifted Child Quarterly, 39*(2), 88–94. https://doi.org/10.1177/001698629503900205

Thiele, T., Singleton, A., Pope, D., & Stanistreet, D. (2016). Predicting students' academic performance based on school and socio-demographic characteristics. *Studies in Higher Education*, *41*(8), 1424-1446. https://doi.org/10.1080/03075079.2014.974528

Ziegler, A. (2005). The actiotope model of giftedness. In R. J. Sternberg, & J. E. Davidson (Eds.), *Conceptions of giftedness* (pp. 411-436). Cambridge, England: Cambridge University Press.

Ziegler, A., & Baker, J. (2013). Talent development as adaption: The role of educational and learning capital. In S. Phillipson, H. Stoeger, & A. Ziegler (Eds), *Exceptionality in East-Asia: Explorations in the actiotope model of giftedness* (pp. 18-39). Routledge.

Ziegler, A., Chandler, K., Vialle, W., & Stoeger, H. (2017). Exogenous and endogenous learning resources in the Actiotope Model of Giftedness and its significance for gifted education. *Journal for the Education of the Gifted, 40*, 310-333. https://doi.org/10.1177/0162353217734376

Ziegler, A., Debatin, T., & Stoeger, H. (2019). Learning resources and talent development from a systemic point of view. *Annals of the New York Academy of Sciences*, 1445, 39—51. https://doi.org/10.1111/nyas.14018

Ziegler, A., & Stoeger, H. (2017). Systemic gifted education. A theoretical introduction. *Gifted Child Quarterly, 61*, 183–193. doi:10.1177/0016986217705 713

Ziegler, A., Stoeger, H., & Balestrini, D. (2017). Systemic gifted education. In J. Riedl Cross, C. O'Rellly & T. Cross (Eds.), *Providing for the special needs of students with gifts and talents* (pp. 15-55). Dublin, Ireland: Kazoo Independent Publishing.

Ziegler, A., Vialle, W., & Wimmer, B. (2013). The actiotope model of giftedness: A short introduction to some central theoretical assumptions. In S. Phillipson, H. Stoeger, & A. Ziegler (Eds.), *Exceptionality in East Asia: Explorations in the actiotope model of giftedness* (pp. 1-17). London, England: Routledge.

Acknowledgements

We wish to thank all the students for their participation in this study. We would also like to thank Patricia Román for her help with back-translation of the scales.

Corresponding author:
Isabel Benítez, PhD
Faculty of Psychology.
Campus de Cartuja s/n. 18071.
Granada (Spain)
Phone: +34 958246234
E-mail: ibenitez@ugr.es

Appendix 1

Original and Spanish versions of the Meaning of Life Questionnaire (MLQ-S)

Original version	Adapted version
1. I understand my life's meaning	Entiendo cuál es el sentido de mi vida
2. I am looking for something that makes my life feel meaningful	Estoy buscando algo que haga que mi vida tenga sentido
3. I am always looking to find my life's purpose	Siempre estoy buscando el propósito de mi vida
4. My life has a clear sense of purpose	Mi vida tiene un claro propósito
5. I have a good sense of what makes my life meaningful	Tengo una idea clara de lo que hace que mi vida tenga sentido
6. I have discovered a satisfying life purpose	He descubierto un propósito vital que me satisface
7. I am always searching for something that makes my life feel significant	Siempre estoy buscando algo que me haga sentir que mi vida tiene un sentido
8. I am seeking a purpose or mission for my life	Busco un propósito o misión para mi vida
9. My life has no clear purpose*	Mi vida no tiene un propósito claro*
10. I am searching for meaning in my life	Estoy buscando sentido a mi vida

*reverse item

Appendix 2

Original and Spanish versions of the Comprehensive Intellectual Humility Scale (CIHS-S)

Original version	Adapted version
1. My ideas are usually better than other people's ideas*	Mis ideas suelen ser mejores que las ideas de otras personas*
2. For the most part, others have more to learn from me than I have to learn from them*	En general, los otros tienen más que aprender de mí que yo de ellos*
3. When I am really confident in a belief, there is very little chance that belief is wrong*	Cuando estoy realmente convencido/a de algo es poco probable que dicha idea sea errónea*
4. I'd rather rely on my own knowledge about most topics than turn to others for expertise*	Prefiero confiar en mi propio conocimiento sobre la mayoría de los temas que recurrir al conocimiento de otros*
5. On important topics, I am not likely to be swayed by the viewpoints of others*	En temas importantes, no es probable que me influyan los puntos de vista de otros*
6. I have at times changed opinions that were important to me, when someone showed me I was wrong	En ocasiones he cambiado opiniones que eran importantes para mí, cuando alguien me ha mostrado que estaba equivocado/a
7. I am willing to change my position on an important issue in the face of good reasons	Estoy dispuesto/a a cambiar mi posición en un tema importante, si hay buenas razones
8. I am open to revising my important beliefs in the face of new information	Estoy abierto/a a revisar creencias importantes para mí si dispongo de información nueva

9. I am willing to change my opinions on the basis of compelling reason	Estoy dispuesto/a a cambiar de opinión ante razones de peso
10. I'm willing to change my mind once it's made up about an important topic	Estoy dispuesto/a a cambiar de opinión sobre un tema importante aunque ya estuviese convencido/a
11. I respect that there are ways of making important decisions that are different from the way I make decisions	Respeto que haya formas de tomar decisiones importantes que sean diferentes a la forma en que yo las tomo
12. Listening to perspectives of others seldom changes my important opinions*	Escuchar las perspectivas de los demás raramente cambia opiniones que son importantes para mí*
13. I welcome different ways of thinking about important topics	Valoro diferentes formas de pensar sobre temas importantes
14. I can have great respect for someone, even when we don't see eye-to-eye on important topics	Puedo sentir un gran respeto por alguien incluso si no coincidimos en temas importantes
15. Even when I disagree with others, I can recognize that they have sound points	Incluso cuando estoy en desacuerdo con otros, puedo reconocer que tiene sentido lo que dicen
16. When someone disagrees with ideas that are important to me, it feels as though I'm being attacked*	Cuando alguien está en desacuerdo con ideas que son importantes para mí, me siento atacado/a*
17. When someone contradicts my most important beliefs, it feels like a personal attack*	Cuando alguien contradice mis creencias más importantes, lo siento como un ataque personal*
18. I tend to feel threatened when others disagree with me on topics that are close to my heart*	Tiendo a sentirme amenazado/a cuando otros están en desacuerdo conmigo en temas que me tocan muy de cerca*
19. I can respect others, even if I disagree with them in important ways	Puedo respetar a los demás, incluso si el desacuerdo con ellos es importante
20. I am willing to hear others out, even if I disagree with them	Estoy dispuesto/a a escuchar a otros, incluso si estoy en desacuerdo con ellos
21. When someone disagrees with ideas that are important to me, it makes me feel insignificant*	Cuando alguien está en desacuerdo con ideas que son importantes para mí, me hace sentir insignificante*
22. I feel small when others disagree with me on topics that are close to my heart*	Me siento poca cosa cuando otros están en desacuerdo conmigo en temas que me tocan muy de cerca

*reverse item.

Appendix 3

Additional items included in the assessment protocol to measure self-criticism in Study 1

Original source	Items
SRS	Es importante para mí plantear una mirada abierta a los otros desde el respeto a su dignidad, sin juzgar sus circunstancias y limitaciones. *It is important for me to look at others with openness and respect for their dignity, without judging their circumstances and limitations (self-criticism 1).*
	Es importante para mí abrirme a los otros y aprender de ellos. *It is important for me to be open to others and learn from them (self-criticism 2).*
	Conozco mis potencialidades y mis limitaciones. *I know my potential and my limitations (self-criticism 3).*
Ad hoc	Soy capaz de autocriticarme y ver mis puntos débiles. *I am able to self-criticize and see my weaknesses (self-criticism 4).*
	Me gusta dialogar con personas que tienen ideas distintas a las mías. *I like to talk to with people with different ideas than I have (self-criticism 5).*
	Valoro positivamente que otras personas puedan expresar libremente sus ideas. *I appreciate the fact that other people can freely express their ideas (self-criticism 6).*
	Creo que exigirme a mí mismo/a me ayuda a mejorar. *I consider that making demands on myself helps me to improve (self-criticism 7).*

NOTE: Direct English direct translation is included for informative purposes

Appendix 4

Spanish revised version of the Meaning of Life Questionnaire (MLQ-S-R)

MLQ-S-R items (items in the previous version of MLQ-S)
1. Entiendo cuál es el sentido de mi vida (MLQ-S-1)
2. Mi vida tiene un claro propósito (MLQ-S-4)
3. Tengo una idea clara de lo que hace que mi vida tenga sentido (MLQ-S-5)
4. He descubierto un propósito vital que me satisface (MLQ-S-6)
5. Mi vida no tiene un propósito claro* (MLQ-S-9)
6. Siempre estoy buscando algo que me haga sentir que mi vida tiene un sentido (MLQ-S-7)
7. Busco un propósito o misión para mi vida (MLQ-S-8)
8. Considero que tengo una misión que orienta mi vida (MLQ-S-10 modified)

*reverse item

Appendix 5

Spanish revised version of the Comprehensive Intellectual Humility Scale (CIHS-S-R)

CIHS-S-R items (items in the previous version of CIHS-S)
1. Soy capaz de autocriticarme y ver mis puntos débiles (self-criticism 4)
2. Me gusta dialogar con personas que tienen ideas distintas a las mías (self-criticism 5)
3. Valoro positivamente que otras personas puedan expresar libremente sus ideas (self-criticism 6)
4. Creo que exigirme a mí mismo/a me ayuda a mejorar (self-criticism 7)
5. Es importante para mí plantear una mirada abierta a los otros desde el respeto a su dignidad, sin juzgar sus circunstancias y limitaciones (self-criticism 1)
6. Es importante para mí abrirme a los otros y aprender de ellos (self-criticism 2)
7. Conozco mis potencialidades y mis limitaciones (self-criticism 3)
8. Mis ideas suelen ser mejores que las ideas de otras personas (CIHS-S-1)
9. En general, los otros tienen más que aprender de mí que yo de ellos (CIHS-S-2)
10. Cuando estoy realmente convencido/a de algo, no hago caso a ideas que me hagan replantearme mi punto de vista* (CIHS-S-3)
11. Prefiero confiar en mi propio conocimiento sobre la mayoría de los temas que recurrir al conocimiento de otros (CIHS-S-4)
12. Cuando un tema es importante para mí, no tengo en cuenta los puntos de vista de otras personas* (CIHS-S-5 modified)
13. En ocasiones he cambiado opiniones que eran importantes para mí, cuando alguien me ha mostrado que estaba equivocado/a (CIHS-S-6)
14. Estoy dispuesto/a a cambiar mi posición en un tema importante, si hay buenas razones (CIHS-S-7)
15. Estoy dispuesto/a a cambiar de opinión ante razones de peso (CIHS-S-9)
16. Respeto que haya formas de tomar decisiones importantes que sean diferentes a la forma en que yo las tomo (CIHS-S-11)
17. Cuando escucho otras perspectivas diferentes a mis opiniones, no suelo tenerlas en cuenta* (CIHS-S-12 modified)
18. Valoro diferentes formas de pensar sobre temas importantes (CIHS-S-13)
19. Puedo sentir un gran respecto por alguien incluso si no coincidimos en temas importantes (CIHS-S-14)
20. Incluso cuando estoy en desacuerdo con otros, puedo reconocer que tiene sentido lo que dicen (CIHS-S-15)
21. Cuando alguien está en desacuerdo con ideas que son importantes para mí, me siento atacado/a (CIHS-S-16)
22. Cuando alguien contradice mis creencias más importantes, lo siento como un ataque personal (CIHS-S-17)
23. Tiendo a sentirme amenazado/a cuando otros están en desacuerdo conmigo en temas que me tocan muy de cerca (CIHS-S-18)
24. Puedo respetar a los demás, incluso si el desacuerdo con ellos es importante (CIHS-S-19)
25. Estoy dispuesto/a a escuchar a otros, incluso si estoy en desacuerdo con ellos (CIHS-S-20)
26. Cuando alguien está en desacuerdo con ideas que son importantes para mí, me hace sentir insignificante (CIHS-S-21)
27. Me siento poca cosa cuando otros están en desacuerdo conmigo en temas que me tocan muy de cerca (CIHS-S-22)

*reverse item

Promoting Equality in Higher Education: Development and Internal Validity of a Selection Test for Science University Degrees in Ecuador

Iván Sandoval[a], Raquel Gilar-Corbí[b], Alejandro Veas[b] & Juan-Luis Castejón[b]

[a] National Polytechnic School of Ecuador. Department of Basic Training.
[b] University of Alicante. Department of Developmental Psychology and Didactics.

Abstract:
The use of tests or examinations for student admission is an extended strategy in most higher education institutions. However, tests in Latin-American countries have not received attention regarding the validity and interpretation of these measures. As such, the aim of this study is to analyse the psychometric properties of a new university entrance examination test in Ecuador in a community sample of 1238 university students (28.10 % female). A two-parameter multidimensional item response theory (MIRT) model was used to calibrate item difficulty and discrimination parameters as well as differential item functioning (DIF) by gender. The final instrument was composed by 71 items, which was considered appropriate to ensure the measurement precision of all levels of students' achievements.

Keywords:
University selection test, polytechnic school, Test Validation, multidimensional item response theory

1 Introduction

The evaluation process is a fundamental tool for training and measuring the impact of educational systems. During the past decades, different studies have attempted to address the quality of measuring external performance assessment tests in international contexts, such as the Programme for International Student Assessment (PISA) and the Trends in International Mathematics and Science (TIMSS), among others (Carnoy et al., 2015; Haladyna & Downing, 2004; Kane, 2013). Moreover, within the fields of science, technology, engineering, and mathematics (STEM), these tests have been considered as major indicators for ranking countries in international assessments (Liou & Bulut, 2020), including education policy initiatives from different countries (Ho, 2016; Lietz & Tobin, 2016; Schmidt & Burroughs, 2016).

The use and application of achievement measures are also crucial in higher education, especially in the tests or examinations designed for student admission and in ensuring the comparability of results (Coe, 2010). In this context, a number of European studies have considered the theoretical and validity measurements of certificate examinations (Baird et al., 2000; Coe 2007; He et al., 2018) or university entrance examinations (Veas et al., 2020a 2020b). However, tests in Latin-American countries have received no attention regarding the validity and interpretation of these types of measures, considering that more social disadvantages exist among students.

This study aims to fill this gap by considering the social and academic selection context of a large polytechnic school in Ecuador to develop and validate a new measurement instrument that assesses the science and language level of students who want to enrol into a STEM degree. This study investigates the test functioning of a new measure under a multidimensional item response theory (MIRT) model, which includes a range of item difficulty and discrimination parameters along a latent construct. Moreover, differential item functioning (DIF) is also explored to determine item gender bias, which may affect the measurement precision of the instrument.

1.1 Social diversity in the higher education selection process

In the Latin-American and Caribbean regions, student retention and dropout are distinct and negative realities in all levels of education. Traditionally, possible reasons focused on the students' characteristics, although these reasons finally moved on from the relation between the students and the institutions to the institutions' responsibility to address a massive and heterogeneous group of students (Braxton et al., 1997; Himmel, 2002).

The negative impact of inequality on student achievement across cultures has proved to be consistent in multiple studies (Alexander et al., 2001; Georges & Pallas, 2010; Ma et al., 2018). The reduction of inequalities in the access to and completion of higher education has been a strategic target in different countries (Eurydice, 2012). This priority is given by the fact that the probability of university enrolment and retention differ substantially across social backgrounds. Important theoretical models have been proposed to provide a consistent explanation of this phenomenon. For instance, Bourdieu and Passeron (1990) referred to the theory of cultural reproduction, where children in the highest classes have

advantages in gaining educational credentials due to their possession of cultural capital. Breen and Golthorpe (1997) conceived of the educational success of students in terms of evaluating costs and benefits and the perceived probability of success outcomes. Another recent view is the consideration of learning capitals from the actiotope model of giftedness (Ziegler & Baker, 2013; Ziegler, Chandler, et al., 2017; Ziegler et al., 2019). This model can be extended to all students' levels of achievement, as it endorses different factors in a dynamic complex system (Ziegler, Balestrini, et al., 2017; Ziegler et al., 2013; Ziegler & Stoeger, 2017). Considering educational capital as external influence of self, inequality may affect to both cultural and social capitals. In this context, it is clear that performance differences in the transition to higher education are also produced by social selection during the earlier stages of schooling.

1.1.1 *Selection process in Ecuadorian higher education institutions*

Special efforts have been taken in Latin-American higher education institutions for them to be considered as relevant actors of social development (Arocena & Sutz, 2005). During the twentieth century, important student movements triggered the so-called University Reform Movement ([URM]: Ribeiro 1971; Tünnermann, 2000), allowing the inclusion of social policies based on the increasing enrolment in higher education despite political or military controversies over the past decades. This spread of democracy in the higher education system has met the goals of stronger teaching and research standards (de Moura Castro & Levy, 2001).

According to social demands, polytechnic schools began to provide qualified professional techniques in fundamental areas of progress in the country. Starting from religious institutions, expert groups determined objective criteria for the development of undergraduate degrees and assessment standards (Contreras & Maluk 2017). For example, polytechnic schools, in comparison to universities, should provide at least 70 % of the professional titles in the basic and applied sciences with the guarantee of excellence and academic rigor among the scientific field. As a more formal example, the National Polytechnic School of Ecuador (NPSE), created on the 30th of August, 1986, is a public higher education institution that is in line with a mass access model policy. As such, and in line with the National Secretariat of Higher Education, Science, Technology, and Innovation (SENESCYT), the student admission process considers vulnerable groups of students from the perspective of social inequality. During this process, an applicant of an Ecuadorian higher education institution must meet certain requirements, such as taking the National Exam of Educational Evaluation (*Ser Bachiller*) and completing the Associated Factors Survey. Although there is no minimum score needed to apply for a degree (it varies depending on the institution), the allocation of places is automatically made according to the application score, the availability of places in each institution, and the demand that exists for a degree in a given period. In short, applicants with the highest scores in the exam are more likely to get a place (SENESCYT 2018). However, in 2014, SENESCYT implemented a positive action policy that would expand access to higher education for socially and economically vulnerable applicants through the Af-

firmative Action Programme. An applicant who has been included in this programme can preferentially apply to 15 % of the academic places offered by public higher education institutions, even if their application score has not met the minimum required for a specific institution during the admission process (Di Caudo, 2015).

The determination of the beneficiaries of the Affirmative Action Programme is achieved through the analysis of each applicant's self-declared information in the Associated Factors Survey. Then, a vulnerability index is calculated, and the lowest values correspond to applicants from traditionally-excluded groups, those who have a disability, or those who are placed in the lowest decile according to their socio-economic status. In this sense, applicants in situations of greater vulnerability are usually assigned to the Affirmative Action population segment.

Since 2017, SENESCYT has included, among the new students entering the levelling course of the NPSE, those from the Affirmative Action population segment. These students' average application scores were observed to be lower than those obtained by other population segments, denoting poor previous academic preparation. During the first year of study, students from vulnerable groups generally show lower academic performance when compared to that of their peers from other population segments. Additionally, during high school, mathematics typically has the lowest indicator as it is perceived as being more complicated.

1.2 Multidimensional item response modelling in test development

Research has begun to apply item response theory models (IRT) to validate measures in different fields (Christensen et al., 2019) with a general use in educational assessment (Embretson, 1984; Hartig & Höhler, 2009), and more specifically, in science education (Kaspersen & Ytterhaug, 2020). Van der Linden (2017) claimed that IRT analysis, which focuses on the quality of items when measuring underlying constructs, perfectly complements classical test theory approaches. In this sense, IRT models has become popular in test construction, including large-scale educational assessment, to optimize item selection and scale validation across diverse populations (Khorramdel & von Davier, 2016). Differing from classical test theory, which considers that an observed test score is composed by a true score and a random component, IRT considers that the probability of a person's expected response to an item is a mathematical function of that person's ability and one or more parameters that characterize the item (Reckase, 2009).

MIRT models have appeared in the research literature since the 1980s (e.g. Bock & Aitken 1981). The purpose of MIRT is to provide a model with an appropriate representation of data (given an incidental vector, θ, which describe the locations of individuals), and structural parameters are used to describe the functioning of the test items in a m-dimensional space, where m is the number of dimensions used to model the data (Reckase, 2009).

MIRT models for dichotomous items (those with two categories) are one of the most important in achievement tests. A multidimensional extension of the two-pa-

rameter logistic model is given by the following equation (Reckase, 2009):

$$P(U_{ij} = 1 | \theta_j, a_i, d_i) = \frac{e^{a_{i\theta' j} + d_i}}{1 + e^{a_{i\theta' j} + d^i}}$$

where U_{ij} is the score for person *j* in item *I*; θ_j is the parameter that describes the ability of the *j*th person in item *I*; and a_i is a parameter related to the maximum slope of the item characteristic curve along the latent construct which measures the item's discriminating power. Given the multidimensional nature of the model, a slope/intercept form, $a\theta + d$, is introduced in the equation, where *d* is the result of the *ab* item interaction (*b* = item difficulty). Therefore, *a* is a $1 \times m$ vector of the item's discrimination parameters and θ is a $1 \times m$ vector of the person's coordinates, with *m* indicating the number of dimensions in the coordinate's space.

The present study

The NPSE has two levelling courses: one for the engineering, sciences, and administrative sciences and one for the superior technological level. The aim of the levelling course, beyond its academic purposes, is to enrol new students to the university context until the study programme's completion. However, according to the information provided by the management and information processes department, approximately 40 % of students who receive a score that is less than or equal to six points in all the subjects of the levelling course (mathematics fundamentals, geometry and trigonometry, physics, chemistry, and language and communication) during the first bimester abandon this course.

The antecedents described are of relevance to the Latin-American science university system, as it is crucial to determine the correct application of selection tests to students who may enrol in the levelling course. Otherwise, possible defects are associated with potential higher education dropout rates. For these reasons, the present study aimed to analyse the psychometric properties of a new selection test under the two-parameter MIRT model. The specific objectives were as follows: (1) to analyse model-data fit of items of the test, (2) measure the precision of items according to individual levels of ability, and (3) invariance properties according to gender.

2 Method

2.1 Participants

The sample comprised of 1238 newly enrolled students (*mean age* = 18. 85, *SD*= 1.84) from the NPSE for the first semester of 2019 (890 males, 348 females). Of these, 955 students (280 males, 675 females) were enrolled in the engineering levelling course, and 288 students (215 males, 68 females) were enrolled in the technology levelling course.

2.2 Measures

The design process of the test began with an initial survey administered to the professors assigned to the levelling courses. These professors identified the elemental topics that students usually present academic difficulties in. These topics were compared to those studied at the higher level of basic general education and the Baccalaureate, and a list of curricular content was elaborated upon to determine students' previous knowledge in order to receive an appropriate score in the levelling courses. Therefore, a diagnosis test was designed following this

content criteria.

The first pilot test composed of 65 multiple choice items that evaluated the following topics: real numbers operations, polynomials operations, factoring system, equations and inequations, and functions in real geometry and trigonometry fields. After its first application in October 2018, the evaluation committee decided that the test should include items on language and communication, with a maximum length of 80 items. Therefore, the instrument was composed of a mathematics section (55 items) and a language and communication section (25 items). All of the items presented four alternative options, of which only one was correct.

2.3 Procedure

The data were collected on the 28th of March, 2019 in paper-pencil form from different classes in the NPSE. Before the application of the test, the university made an official announcement to students via email and provided an instructive link in the university webpage. Furthermore, professors in the department of basic training were recruited for responsibility over controlling the students during the test application as well as distributing and collecting the material. The duration of the test was two hours in accordance with the evaluation criteria.

Students' responses were sent to the admission and registration unit for computer correction. Every correct response was scored as 1 point, whereas incorrect responses were scored as 0 points. Incorrect responses were not penalised. Punctuation in each section was computed as 50 % of the total score.

2.4 Data analysis

Data analysis comprised of two phases of validation. First, the fitting quality of each item considering the multidimensional structure was analysed through the expected values of infit (weighted) and outfit (unweighted) mean square error, and statistics were determined to be between -2 and +2 according to standard criteria (Fan 1998). Second, regarding the content, this study employed the fitting quality of each item considering the multidimensional structure of the instrument. Next, with respect to the generalisability aspect of validity, this study conducted differential item functioning (DIF: Holland & Wainer, 1993) analysis among gender. A difference of 0.5 logits in the overall item difficulty across groups was considered as a substantial DIF. The mean item parameters were set to be equal over groups so that the differences in the parameter estimates could be directly compared. Moreover, the item discrimination index was analysed and a good index criterion was considered to be above 0 and below 2 (De Ayala, 2009). The parameters were estimated using the computer programme Conquest Version 2 (Wu et al., 2007) via the maximum likelihood method.

3 Results

The two dimensions of the entrance examination test were calibrated simultaneously. Table 1 shows the difficulty estimates, fit statistics, discrimination estimates, and DIF magnitudes for each item. All of the items showed excellent infit and outfit values, and most of them were close to 1.00. The discrimination parameters showed acceptable values in all of the items that belonged to the mathematics subscale. With

respect to the language and communication subscale, items 56 (-0.33), 57 (-0.20), 64 (-0.07), 74 (-0.07), 75 (-0.04), 76 (-0.16), and 77 (-0.15) showed values below 0. Therefore, these items did not have enough power to discriminate between the more and less able students on the language and communication subscale.

DIF analyses were conducted to assess the model-data fit across gender. As indicated in Table 1, items 31, 56, and 61 showed significant DIF, implying that these three items were more difficult for females. The largest values were for items 56 and 61, which were included in the language and communication subscale.

When applying the purification procedure (Lord, 1980), items with non-adequate discrimination parameters or no DIF values were removed, and a new item parameter was implemented which considered 71 items: 54 belonged to the mathematics subscale and 17 belonged to the language and communication subscale.

An item-person map is provided in Figure 1, as it is possible to calibrate a person's measurement from low to high and item difficulty from easy to hard along the same latent trait scale. The two continuums on the left side of the figure indicate students' measures in the two dimensions of the test. Individuals who had high scores are placed at the top of the continuum and those who had lower scores are placed at the bottom. Moreover, the items that fall into each of the two dimensions are clustered on the right side. All of the items are distributed reasonably well along the latent construct. The students located at the medium side of the scale were targeted by the majority of items. The most difficult items were 59, 60, and 64, which belonged to the language and communication subscale; the easiest items were 7, 42, and 36, which belonged to the mathematics subscale. These items targeted an important proportion of low-ability students.

4 Discussion

This study aimed to analyse the psychometric properties of a newly developed version of a university entrance examination test in a sample of university students enrolled in the ESPN, one of the largest public institutions in Ecuador. This instrument was intended to ensure that students of a minimum curriculum level had access to the levelling courses under a global access policy that focuses on the population's social diversity and vulnerability.

To gain a deeper understanding of the measurement precision, a two-parameter MIRT was implemented. Considering the initial 80 items distributed in two subscales (mathematics and language and communication), the results showed excellent item fit values. Nine items showed poor discrimination parameters or DIF; without these items, the new estimation provided acceptable values for all parameters. In general terms, the mathematic subscale showed better parameter values than the language and communication subscale. The item-person map showed that item difficulty was reasonably spread at the top of the map, as the main objective of the scale was to detect all students' achievement levels using adequate measurement precision.

By using construct validation measures, it is possible to extend appropriate measurement practices in Latin-American universities and ensure equality processes through innovative network policies among institutions (Arocena & Sutz, 2001). Because of the widespread concern over the social needs of

Items parameters

Items	Item difficulty (SE)	Infit	Outfit	Item discrimination	Gender DIF
1	0.18(0.07)	1.00	1.00	0.89	0.21
2	0.29(0.08)	0.99	1.04	1.29	0.15
3	0.45(0.07)	0.99	0.99	1.17	0.09
4	0.76(0.07)	0.99	1.01	0.75	0.04
5	-0.24(0.07)	1.00	1.01	0.95	0.18
6	-0.04(0.08)	0.99	1.02	1.47	0.00
7	-1.61(0.12)	1.00	1.01	1.85	0.03
8	-0.36(0.07)	1.00	0.98	1.02	0.29
9	-0.28(0.07)	0.99	0.97	1.03	0.06
10	0.63(0.09)	0.99	1.07	1.88	0.13
11	0.36(0.06)	1.00	1.00	0.48	0.05
12	1.03(0.07)	0.99	1.02	0.82	0.06
13	1.29(0.08)	0.99	1.03	0.98	0.10
14	-1.18(0.08)	1.01	0.98	0.81	0.22
15	0.53(0.07)	0.99	1.01	0.99	0.10
16	-0.28(0.06)	1.00	0.99	0.62	0.07
17	-0.16(0.07)	0.99	1.02	1.19	0.05
18	-0.42(0.08)	0.99	0.97	1.63	0.13
19	0.86(0.08)	0.98	1.02	1.23	0.20
20	0.43(0.07)	0.99	1.00	1.04	0.00
21	-0.17(0.06)	1.00	1.00	0.52	0.22
22	1.65(0.09)	0.99	1.05	0.94	0.01
23	-0.31(0.08)	1.00	0.99	1.14	0.02
24	-0.52(0.08)	0.99	1.05	1.40	0.03
25	-0.65(0.08)	0.99	1.12	1.41	0.03
26	-0.47(0.07)	0.99	1.05	1.04	0.02
27	-0.65(0.07)	1.00	1.00	0.96	0.28
28	-0.59(0.07)	0.99	1.01	1.10	0.08
29	0.95(0.07)	0.99	1.02	0.72	0.20
30	-0.44(0.08)	0.98	1.09	1.49	0.04
31	0.13(0.07)	0.99	1.01	1.07	0.23
32	1.01(0.07)	0.99	1.02	0.75	0.21
33	0.38(0.06)	1.00	1.01	0.64	0.21
34	0.72(0.08)	0.98	0.99	0.85	0.19
35	-0.40(0.07)	0.99	1.04	1.43	0.29
36	-1.35(0.1)	1.01	0.97	1.43	0.26
37	-0.80(0.07)	1.00	0.99	0.63	0.26
38	0.32(0.08)	0.99	0.99	1.36	0.14
39	0.55(0.08)	0.99	1.05	1.31	0.54**
40	0.53(0.09)	0.99	1.03	1.80	0.14
41	0.62(0.07)	0.99	1.01	0.96	0.12
42	-1.49(0.09)	1.01	1.01	1.00	0.20
43	-0.64(0.07)	1.01	1.00	0.71	0.22
44	0.95(0.07)	0.99	1.00	0.61	0.06
45	-0.77(0.08)	1.01	0.97	0.89	0.05
46	0.19(0.07)	0.99	1.00	1.03	0.22
47	0.68(0.08)	0.98	1.02	0.96	0.18
48	0.28(0.07)	0.99	1.01	0.92	0.13
49	0.47(0.07)	1.00	1.00	0.47	0.07
50	0.83(0.07)	0.99	1.01	0.47	0.12
51	1.28(0.07)	0.99	1.01	0.48	0.11
52	1.39(0.08)	0.99	1.02	0.55	0.24
53	0.23(0.08)	0.99	0.98	1.63	0.11
54	0.23(0.07)	1.00	0.98	0.82	0.32
55	1.36(0.07)	0.99	1.01	0.34	0.20
56	0.92(0.07)	1.00	1.00	-0.33*	0.71**
57	-0.07(0.06)	1.00	1.00	-0.20*	0.21
58	0.35(0.07)	1.00	1.00	0.81	0.32
59	2.12(0.1)	1.00	1.00	0.50	0.00
60	2.07(0.09)	1.00	1.00	0.22	0.11
61	0.30(0.06)	1.00	1.00	0.12	0.70**
62	1.19(0.07)	1.00	1.00	0.18	0.47
63	-0.34(0.07)	1.00	1.00	1.01	0.22
64	2.02(0.09)	1.00	1.00	-0.07*	0.07
65	1.73(0.08)	1.00	1.00	0.09	0.07
66	1.00(0.07)	1.00	1.00	0.54	0.48
67	-0.03(0.06)	1.00	1.00	0.12	0.06
68	-0.15(0.06)	1.00	1.00	0.48	0.13
69	0.88(0.07)	1.00	1.01	0.41	0.32
70	1.18(0.07)	1.00	1.00	0.42	0.16
71	1.51(0.10)	1.00	1.04	1.17	0.17
72	0.49(0.06)	1.00	1.01	0.44	0.03
73	1.36(0.07)	1.00	1.00	0.09	0.19
74	1.42(0.07)	1.00	1.00	-0.07*	0.06
75	0.09(0.05)	1.00	1.00	-0.04*	0.08
76	1.40(0.07)	1.00	1.00	-0.16*	0.01
77	1.67(0.08)	1.00	1.00	-0.15*	0.22
78	0.09(0.06)	1.00	1.00	0.49	0.29
79	1.77(0.08)	1.00	1.00	0.08	0.14
80	-0.51(0.06)	1.00	1.00	0.51	0.02

Note. *SE* = Standard error; DIF = Differential Item Functioning; * = poor discrimination value; ** = substantial DIF.

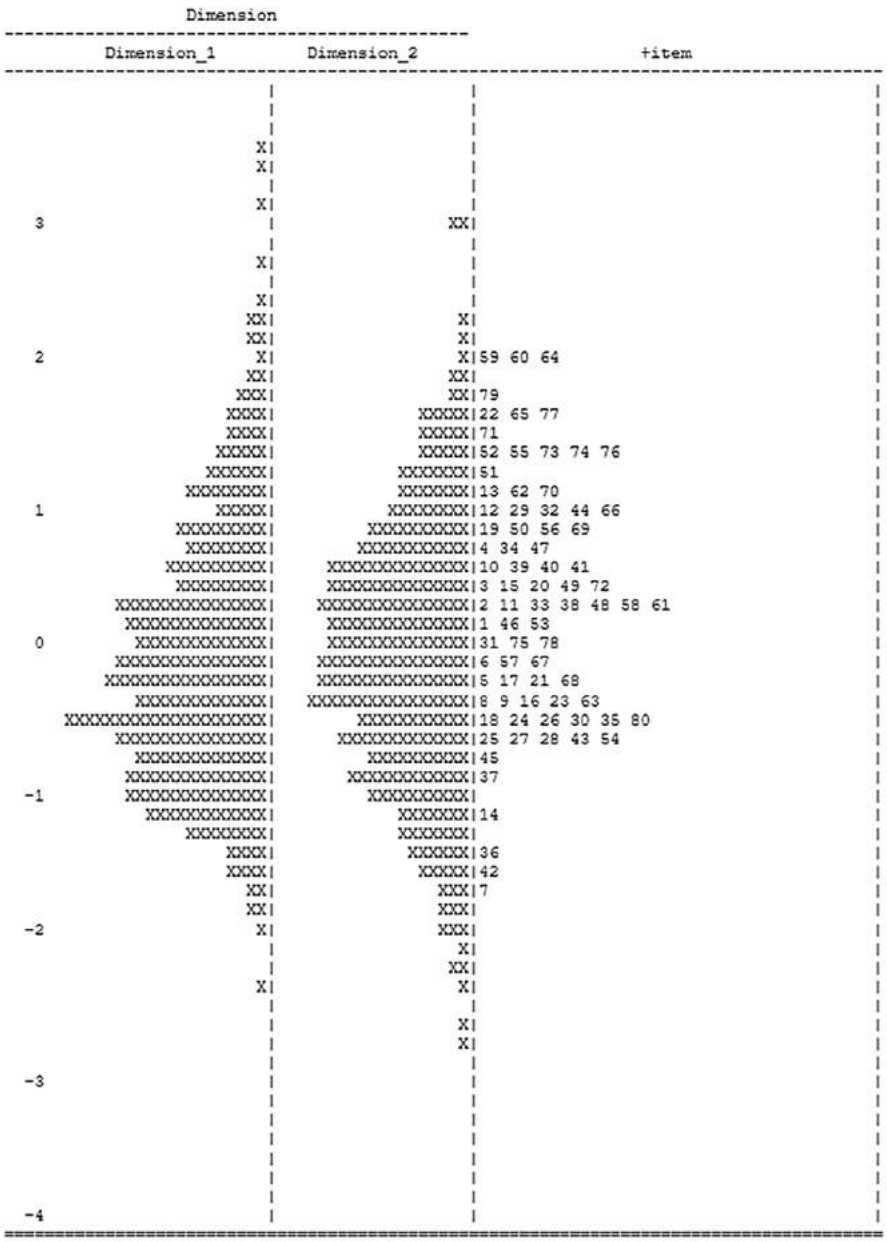


Figure 1 Map of latent distributions and responde model parameters.

Note. Each 'X' represents 4.4 cases

higher education systems, academic quality seems to be increasing. Hence, new public policies for student recruitment and selection should be able to obtain stronger support than in past years.

The application of common standard criteria in examination tests considers the variability of social background differentials in the enrolment and retention probabilities across student profiles (Jordan et al., 1996). This possibility allows improved actions in the environment based on a depth analysis of learning capitals (Ziegler & Baker, 2013). The main advantage is to establish more objective decisions regarding academic and professional trajectories that does not depend on the possible indirect cost of education or on less prestigious choices due to being more risk averse. In terms of episodic learning capital, resource investments imply that better-quality education and better involvement fosters cognitive and non-cognitive skills (Carneiro & Heckman, 2002; Ziegler, 2005). Latin-American institutions might decide on how to collaborate with national development using divergent strategies and consolidate an open policy without being worried about disadvantages, such us losing the effectiveness of the application of curriculum and knowledge through standard measurement. To this end, active purposes are intended to analyse the pragmatic and contextual approach to examining the process, the results obtained, and the methods used by various organisations (Sondergeld & Koskey, 2011) to ensure the principles of equity and equal opportunity for university admissions.

In conclusion, this study initiates an effective analysis in Ecuador that analyses test scores using advanced psychometric methods such as MIRT as an extension of European studies and American studies on official certificate examinations. However, it is important to bear in mind certain limitations, which may guide future research on this topic. First, it should be noted that the data used herein were students enrolled only in ESPN. Larger samples from other Ecuadorian universities may enable both better estimates of the achievement measures of students and deeper comparisons between the rates of access to levelling courses between the institutions. In this specific context, the use of MIRT models enable comparisons to determine appropriate measures for equity. Second, future analyses should include possible influences of the individual selection of science subjects or the effects of educational reforms on testing (Hübner et al., 2019; Korobko et al., 2008).

5 References

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2001). Schools, achievement, and inequality: A seasonal perspective. *Educational Evaluation and Policy Analysis*, 23(2), 171-191. https://doi.org/10.3102/06123737023002171

Arocena, R. & Sutz, J. (2001). Changing knowledge production and Latin American universities. *Research Policy, 30*(8), 1221-1234. https://doi.org/10.1016/S0048-7333(00)00143-8

Arocena, R., & Sutz, J. (2005). Latin American Universities: From an original revolution to an uncertain transition. *Higher Education, 50*, 573-592. https://doi.org/10.1007/s10734-004-6367-8

Baird, J., Cresswell, M., & Newton, P. (2000). Would the real gold standard please step forward? *Research Papers in Education, 15*(2), 213-229. https://doi.org/10.1080/026715200402506

Bordieu, P., & Passeron, J.C. (1990). *Reproduction in education, society, and culture*. Sage.

Braxton, J., Sullivan, A., & Johnson, R. (1997). Appraising Tinto's theory of college student departure. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research* (pp. 107-164). Springer.

Breen, R., & Glothorpe, J. (1997). Explaining educational differentials: towards a formal rational action theory. *Rationality and Society, 9*, 275-305.

Carneiro, P., & Heckman, J. (2002). The evidence on credit constraints in postsecondary schooling. *Economic Journal, 112*, 705-734.

Carnoy, M., Khavenson, T., & Ivanova, A. (2015). Using TIMSS and PISA results to inform educational policy: a study of Russia and its neighbours. *Compare, 45*(2), 248-271.

Christensen, K. S., Oernboel, E., Nielsen, M. G., & Bech, P. (2019). Diagnosing depression in primary care: A Rasch analysis of the major depression inventory. *Scandinavian Journal of Primary Health Care,* 37, 105-112. https://doi.org/10.1080/02813432.2019.1608039

Coe, R. (2007). Common examinee methods for monitoring the comparability of examination standards. In P. Newton, J. Baird, H. Goldstein, H. Patrick and P. Tymms (Eds.), *Techniques for monitoring the comparability of examination standards.* London, UK: Qualifications and Curriculum Authority.

Coe, R. (2010). Understanding comparability of examination standards. *Research Papers in Education, 25*(3), 271-284. https://doi.org/10.1080/02671522.2010.498143

Contreras, F. A., & Maluk, S. A. (2017). Análisis descriptivo del gobierno universitario ecuatoriano: una mirada desde los cambios legislativos [descriptive análisis of the Ecuadorian university government: an análisis from the legislative changes]. *Revista Electrónica de Investigación educativa, 19*(2), 22-37. https://doi.org/10.24320/redie.2017.19.2.866

De Ayala, R. J. (2009). *The theory and practice of Item Response Theory*. The Guilford Press.

de Moura Castro, C., & Levy, D. C. (2001). Putting reality ahead of myths: A key to reform in Latin America. *International Higher Education, 22*, 16-18.

Di Caudo, M. V. (2015). Política de cuotas en Ecuador: me gané una beca para estudiar en la universidad. *Ponto-e-Vírgula, 17*, 196-218.

Embretson, S. E. (1984). *A general latent trait model for responses processes.* Psychometrika, 49, 175-186.

Eurydice (2012). *Key data on education in Europe 2012*. European Commission.

Fan, X. (1998). Item Response Theory and Classical Test theory: an Empirical Comparison of their Item/person statistics. *Educational and Psychological Measurement, 58*(3), 357-381. https://doi.org/10.1177/0013164498058003001

Georges, A., & Pallas, A. M. (2010). New look at a persistent problem: Inequality, Mathematics achievement, and teaching. The Journal of Educational Research, 103(4), 274-290. https://doi.org/1080/00220670903382996

Haladyna, T. M. & Downing, S. M. (2004). Construct-irrelevant variance in high-stakes testing. *Educational Measurement: Issues and Practice, 23*(1), 17-27. https://doi.org/10.1111/j.1745-3992.2004.tb00149.x

Hartig, J., & Höhler, J. (2009). Multidimensional IRT models for the assessment of competencies. *Studies in Educational Evaluation, 35*(2-3), 57-63. https://doi.org/10.1016/j.stueduc.2009.10.002

He, Q., Stockford, I., & Meadows, M. (2018). Inter-subject comparability of examination standards in GCSE and GCE in England. *Oxford Review of Education, 44*(4), 494-513. https://doi.org/10.1080/03054985.2018.1430562

Himmel, E. (2002). Modelos de análisis de la deserción estudiantil en la educación superior [models of analysis of student's dropout in Higher Education]. *Calidad de la Educación, 17*, 91-108.

Ho, E. S. C. (2016). The use of large-scale assessment (PISA): insight for policy and practice in the case of Hong Kong. *Research Papers in Education, 31*(5), 516-528. https://doi.org/10.1080/02671522.2016.1225351

Holland, P. W., & Wainer, H. (Eds.). (1993). *Differential item functioning*. Hillsdale, NJ: Erlbaum.

Hübner, N., Wagner, W., Hochweber, J., Neumann, M., & Nagengast, B. (2019). Comparing apples and oranges: Curricular intensification reforms can change the meaning of students' grades! *Journal of Educational Psychology.* https://doi.org/10.1037/edu0000351

Jordan, W. L., Lara, J., & Mc Partland, J. M. (1996). Exploring the causes of early dropout among race-ethnic and gender groups. *Youth and Society, 28*(1), 62-94.

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement, 50*(1), 1-73. https://doi.org/10.1111/jedm.12000

Kaspersen, E., & Ytterhaug, B. O. (2020). Measuring mathematical identity in lower secondary school. *International Journal of Educational Research, 103,* 101620. https://doi.org/10.1016/j.ijer.2020.101620

Khorramdel, L., & von Davier, M. (2016). Item Response Theory as a Framework for test construction. in K. Schweizer & C. DiStefano (Eds.). *Principles and methods of test constructions* (pp. 52-82). Göttingen: Hogrefe.

Korobko, O. B., Glas, C. A., Boskeer, R. J., & Luyten, J. W. (2008). Comparing the difficulty of examination subjects with item response theory. *Journal of Educational Measurement, 45*(2), 139-157. https://doi.org/10.1111/j.1745-3984.2007.00057.x

Lietz, P., & Tobin, M. (2016). The impact of large-scale assessment in education on education policy: evidence from around the world. *Research Papers in Education, 31*(5), 499-501. https://doi.org/10.1080/0267122.2016.1225918

Liou, P. Y., & Bulut, O. (2020). The effects of item format and cognitive domain on students' science performance in TIMSS 2011. *Research in Science Education, 50*, 99-121. https://doi.org/10.1007/s11165-017-9682-7

Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Lawrence Erlbaum.

Ma, Y., Hou, X., Huang, J., Wang, W. Li, Y., Zhou, X., & Du, X. (2018). Educational inequality and achievement disparity: An empirical study of migrant children in China. *Children and Youth Services Review, 87,* 145-153. https://doi.org/10.1016/j.childyouth.2018.02.026

Reckase, M. D. (2009). *Multidimensional Item Response Theory*. Springer.

Ribeiro, D. (1971). *La Universidad Latinoamericana [The Latin-American University]*. Caracas: Ediciones de la Biblioteca de la Universidad Central de Venezuela.

SENESCYT (2018). *Proceso de Admisión 2018 [Admission Process 2018]*. Available at: https://admision.senescyt.gob.ec/soluciones/existe-puntaje-minimo-ingresas-una-carrera/.

Schmidt, W. H., & Burroughs, N. A. (2016). Influencing public school policy in the United States: the role of large-scale assessments. *Research Papers in Education, 31*(5), 567-577. https://doi.org/10.1080/02671522.2016.1225355

Sondergeld, T., & Koskey, K. (2011). Evaluating the impact of an urban comprehensive school reform: An illustration of the need for mixed methods. *Studies in Educational Evaluation, 37*, 91-107. https://doi.org/10.1016/j.stueduc.2011.08.001

Tünnermann, C. (2000). *Universidad y Sociedad. Balance histórico y perspectivas desde Latinoamérica [University and society. Historical balance and perspectives from Latin-America]*. Caracas: Universidad Central de Venezuela.

Van der Linden, W. J. (2017). *Handbook of item response theory, volume three: Applications*. Chapman and Hall/CRC. https://doi.org/10.1201/9781315117430

Veas, A., Benítez, I., Navas, L., & Gilar-Corbí, R. (2020a). A comparative analysis of university entrance examinations using the construct comparability approach. *Revista de Educación, 388*, 65-83. https://doi.org/10.4438/1988-592X-RE-2020-388-447

Veas, A., Navas, L., Pozo-Rico, T., & Miñano, P. (2020b). University Entrance Examinations in Spain: using the construct comparability approach to analyze standards quality. *Frontiers in Psychology, 11*: 127. https://doi.org/10.3389/fpsyg.2020.00127

Wang, W. C., Yao, G., Tsai, Y. J., Wang, J. D., & Hsich, L. (2006). Validating, improving reliability, and estimating correlation of the four subscales in the WHOQOL-BREF using multidimensional Rasch analysis. *Quality of Life Research, 15*, 607-620. https://doi.org/10.1007/s11136-005-4365-7

Wu, M. L., Adams, R. J., Wilson, M. R., & Haldane, S. A. (2007). *ACER Conquest, version 2.0: Generalized item response modelling software.* Camberwell: Australian Council for Educational Research.

Ziegler, A., & Baker, J. (2013). Talent development as adaption: The rôle of educational and learning capital. In S. Phillipson, H. Stoeger, & A. Ziegler (Eds.), *Exceptionality in East-Asia: Explorations in the actiotope model of giftedness* (pp.18-39). Routledge.

Ziegler, A., Chandler, K., Vialle, W., & Stoeger, H. (2017). Exogenous and endogenous learning resources in the Actiotope Model of Giftedness and its significance for gifted education. *Journal for the Education of the Gifted, 40*, 310-333. https://doi.org/10.1177/0162353217734376

Ziegler, A., Debatin, T., & Stoeger, H. (2019). Learning resources and talent development from a systemic point of view. *Annals of the New York Academy of Sciences*, 1445, 39—51. https://doi.org/10.1111/nyas.14018

Ziegler, A., & Stoeger, H. (2017). Systemic gifted education. A theoretical introduction. *Gifted Child Quarterly, 61*, 183–193. doi:10.1177/0016986217705 713

Ziegler, A., Stoeger, H., & Balestrini, D. (2017). Systemic gifted education. In J. Riedl Cross, C. O'Rellly & T. Cross (Eds.), *Providing for the special needs of students with gifts and talents* (pp. 15-55). Dublin, Ireland: Kazoo Independent Publishing.

Ziegler, A., Vialle, W., & Wimmer, B. (2013). The actiotope model of giftedness: A short introduction to some central theoretical assumptions. In S. Phillipson, H. Stoeger, & A. Ziegler (Eds.), *Exceptionality in East Asia: Explorations in the actiotope model of giftedness* (pp. 1-17). London, England: Routledge. Ziegler, A., Vialle, W., & Wimmer, B. (2013). The actiotope model of giftedness: A short introduction to some central theoretical assumptions. In S. Phillipson, H. Stoeger, & A. Ziegler (Eds.), *Exceptionality in East Asia: Explorations in the actiotope model of giftedness* (pp. 1-17). London, England: Routledge.

Declarations

Funding
Not applicable.

Conflicts of interests
The authors declare that there is not any conflict of interest.

Authors' contribution
Iván Sandoval: Theoretical review, data collection, manuscript review.
Raquel Gilar-Corbí: Theoretical Review, data collection.
Alejandro Veas: Theoretical review, data analyses.
Juan Luis Castejón: manuscript review.

Corresponding author:
Alejandro Veas, PhD
University of Alicante
Department of Developmental Psychology and Didactics
PO 03690, San Vicente del Raspeig
Alicante, Spain
E-mail: Alejandro.veas@ua.es

Educational and Learning Resources in a Greek Student Sample: QELC factor structure and methodological considerations

Aikaterini D. Gari, Kostas Mylonas, Vassiliki Nikolopoulou,Irina Mrvoljak

Department of Psychology, National and Kapodistrian University of Athens

Abstract:
This research, inspired by the Actiotope Model of Giftedness as a holistic approach to gifted students' potential, aims to address theoretical, methodological, and structural aspects of the Questionnaire of Educational and Learning Capital (QELC) in a Greek sample of children and young teenagers (10- to 14-year-olds). QELC (Ziegler & Baker, 2013) was administered in classrooms, assessing the students' educational and learning resources as two forms of capital, the educational one (5 subscales) and the learning capital (5 subscales) as self-reported by students. The total sample consisted of 740 students, from 16 Greek primary and secondary public schools in Athens, other cities and some non-urban areas of Greece. Cronbach's α indices for the Educational capital and Learning capital subscales were satisfactory ranging from .74 to .85. Confirmatory Factor Analysis, through 1st and 2nd order models, confirmed the theoretical structure of 10 distinct subscales and also supported to a satisfactory extent two higher-order factors. Statistically significant differences were observed for the educational didactic scale and the attentional learning scale by place of residence, and by the two age-bands of students (10- to 11- and 12- to 14-year-olds). The results are discussed on the basis of the QELC verified factor structure and its theoretical implications, along with its cross-cultural perspective.

Keywords:
educational and learning capital, Actiotope Model of Giftedness, first and second-order factor-structure, cross-cultural perspective, the "appropriate coefficients" question

Introduction

As early as the '90s, it has been argued that excellence appears through the individual-environment interaction, rather than through individual traits (Csikszentmihalyi, 1988, 1996). Since then, 30 years have passed but in the twenty-first century, it is considered a necessity "to abandon g-centrism in identifying gifted individuals in the realm of successful intelligence, which is more closely related to practical, adaptive skills" (Tannenbaum, 2000, p. 50). From a social-psychological approach, giftedness was identified by Tannenbaum (1983, 1988) with great performance or high productivity, through the combination of five factors: superior general intelligence, exceptional special aptitudes, non-intellective facilitators, environmental influences – on social, physical and intellectual levels, and situational factors of chance or luck in life stages. All these five factors have to be combined at a high level in order to connect potential with high-level accomplishment; therefore, success depends upon a combination of all five facilitators, and failure can result from even a single deficit.

Exceptional actions are likely to be achieved within a highly stimulating environment within family, school and peer groups, which may facilitate exceptional use of excellent characteristics and also overcome effectively all deficits and obstacles (Mönks, Heller, & Passow, 2000; Mönks, Van Boxtel, Roelefs, & Sanders, 1986). However, as regards school environment and gifted student identification and educational provision, adequate motivation and support are intended to be implemented *after* the identification of the gifted student, an identification mostly based on traits of giftedness and psychometric processes (Gari, Kalantzi-Azizi, & Mylonas, 2000; Phillipson, Phillipson, & Eyre, 2011). However, from a systemic point of view, high learning achievements and effective fulfillment of the student's potential remain at risk. Psychosocial support and educational provisions often depend on students' specific traits and their assessment, along with the possible coordinated efforts of students towards a set of feasible goals and aims; therefore, the benefits are likely to be temporary and restricted by the assessed traits themselves and the relevant goals set (Ziegler & Phillipson, 2012).

For the Actiotope Model of Giftedness (AMG), the focus is on the interaction between a rich stimulating school community and the extreme inner potential of students, which can effectively create excellent achievement. Instead of analyzing students' traits, the unit of analysis of giftedness and any kind of excellence is the *actiotope*, which is a combination of variables within the individual and his/her interactions in the environment in which he/she acts – material, social, and informational (Ziegler, 2005). The term "actiotope" is derived from two Greek words, the verb *ἄγω* that means "to drive towards" and the word *τόπος* that means "place". A student, especially a gifted one, becomes an "ἄκτωρ"", that is, a person who moves towards a direction that bears a specific meaning. Thus, an actiotope focuses on an individual's acting and interacting with contextual ecological, biological, and social levels of a community system, containing all the specific individual qualities that are unique, viewed from the specific environmental context for each individual (Ziegler & Stoeger, 2008, 2017; Ziegler, Stoeger, Harder, Park, Portešová, & Porath, 2014). Therefore, the AMG is interested in a holistic approach to students' potential,

founded on students' goal-directed actions and their specific meanings, as well as their interdependence with other school community members actions, along with general acceptance and support on a material, cultural and psychosocial basis. High abilities, extraordinary achievement and excellent development of skills are perceived by the AMG as the best adaptation to the environmental demands and the best response to the available inner resources.

Under the systemic principle that a system needs effective and adequate resources in order to evolve and function as a unit (Bateson, 1979; von Bertalanffy, 1968), external and internal resources seem to highly regulate the gradual fulfillment of an individual-system potential. Beyond the linear cause-and-effect approach to giftedness, excellence can be achieved as a result of a dynamic combination of excellent traits and facilitating parameters, but it can also become a cause of developing new skills, adopting new attitudes and values, and exploring alternative meanings of action. On the basis of the equifinality systemic principle, excellent outputs may be achieved from a great variety of starting points, and the opposite, different results are likely to appear despite the fact that the starting point is absolutely the same (Ziegler & Phillipson 2012). Therefore, whatever can nourish and motivate excellent development, either as a fruitful input of the inner system itself or derived by the environment, can actually improve high abilities, expert skills and eminent actions urging gradually towards more future effective system outputs (Bateson, 1979; Dowling, 1985; Molnar, 1986; Ziegler, 2005; Ziegler & Baker, 2013; Ziegler & Phillipson, 2012; Ziegler & Stoeger, 2017; Ziegler, Vialle, & Wimmer 2013).

Resources for students may be derived either from the environment, *as exogenous resources,* or from the inner self, as a set of *endogenous resources* that may facilitate, support and strengthen learning achievement. Adequate exogenous resources may facilitate students' action repertoires effectively and urge their potential internal resources to high levels of fulfillment. The former type of resources refers to educational resources mostly derived by the school community members actions and educational processes, e.g. parents, instructors, peer group etc., on a psychosocial, financial and material level; the latter type refers to inner resources, localized in students themselves that regulate learning processes oriented towards what each student, as a unique entity in a specific environmental setting, has access to. These exogenous and endogenous resources interact dynamically and co-construct two different types of capital for students' lives within the school community: *the educational capital*, on economic, cultural, social, materialistic (infrastructural) and educational/didactic levels, and *the learning capital*, on biological/physiological grounds, on selecting optimal actions in order to satisfy demanding needs and desired goals, on performing goals and aims, on experience, and on attentional focus to what is crucial or the most important.

Therefore, ten distinct forms of capital are available to each student; on the educational level: economic, cultural, social, infrastructural and didactic capitals; and on the learning level: organismic, actional, telic, episodic and attentional capitals (Vladut, Vialle, & Ziegler, 2015; Ziegler & Baker, 2013; Ziegler, Balestrini, & Stoeger, 2018; Ziegler, Chandler, Vialle, & Stoeger, 2017; Ziegler & Phillipson, 2012). *Economic educational capital* refers to material valuables

and every kind of valuable goods or procedures that may support learning and education (Ziegler & Baker, 2013). For example, the hard economic crisis in Greece has seriously restricted economic and scientific resources on all levels of the educational system; approximately 500,000 adults aged 40 or less, holding at least B.Sc. degrees (e.g. physicians, lawyers, engineers, dentists, academics, etc.) have migrated to other European countries, the United States, Australia and Asian countries (European Commission, 2017; Gari, 2019) in search of better working and living conditions; *Cultural educational capital* includes values, attitudes, ideologies, ideal symbols and ways of thinking that may facilitate or set obstacles to learning processes. In a study on 568 Greek teachers of both genders (50.4 % females), approximately half of whom worked at state schools in Athens in both primary (56.3 %) and secondary education (43.7 %), teachers' attitudes were found to be positive towards enriching educational processes for the gifted students, the social value of gifted students, and the idea of giftedness as a social capital (Gari, 2016); *Social capital* includes all educational institutions, scholars and other respectful individuals and organizations that contribute to the effectiveness of educational and learning procedures; *Infrastructural educational capital* refers to all goods and materials available (buildings, classrooms, libraries, alternative methods of study etc.) to facilitate educational and learning procedures or – if absent – the opposite, to hinder them; *the didactic educational capital* includes experts, teachers, trained instructors, programs and curricula, included in the implementation of educational processes (Ziegler & Baker, 2013). *Organismic learning capital* regards physical and mental health, along with physiological strengths vs. deficits; *Actional learning capital* includes all sets and patterns of individuals' actions, the "action repertoire of a person" (Ziegler & Baker, 2013, p. 30); *Telic learning capital,* a term that derives from the Greek word "τέλος", which means *moving towards an end or an ultimate point,* includes all goals and aims that an individual sets in order to create chances for meeting his/her needs; *Episodic learning capital,* derived from the Greek word "επεισοδιακός" which means "to be based on a set of prior experiences", in order to select the optimal actions for achieving goals and aims of important meaning, within the current situation; finally, *attentional learning capital* refers to the levels of attention effectiveness towards what is important or of crucial significance for each specific circumstance, on both quantitative and qualitative levels (Ziegler & Baker, 2013).

The empirical basis of the educational and learning capital forms was also extended through a recent study which separately examined three different samples – 365 primary school students in the United Arab Emirates, 90 German female STEM professionals, and 74 German long-distance runners. For the sample of primary school students, it was shown that, beyond IQ, QELC scores predict excellence in academic achievement. For the group of professionals, the adequate availability of exogenous and endogenous resources seemed to play an important role in higher skill development, excellence in performance and a more effective professional development. In addition, the process itself of increasing achievement level seemed to contribute to the gradual possession of more and more exogenous and endogenous resources and their better use, forming "a virtuous circularity" of outcomes (Ziegler, Debatin, & Stoeger, 2019).

Throughout the world the concepts of giftedness and excellence vary. For example, the Western vs. the East Asian conceptions of giftedness focus on individualistic vs. collectivistic social perceptions and the mind-body dualism of the Western Enlightenment vs. the Confucian holistic outlook, respectively. However, in Western research on giftedness a conceptual bias seems to be apparent, while the necessity of cross-cultural research on giftedness remains a demand, due to the small number of cross-cultural studies conducted, mostly on the topics of conceptions of giftedness, identification strategies and educational provision (Stoeger, Balestrini, & Ziegler, 2018). For the QELC model of ten forms of capital, a first set of cross-cultural comparisons was attempted in a study with students in China, Germany and Turkey (mean age ranging from 12.70 to 13.98 years), which effectively supported the QELC psychometrically, and with respect to its construct and concurrent validity (Vladut, Liu, Leana-Taşcilar, Vialle, & Ziegler, 2013). Later, QELC was successfully supported in Israel as well (Paz-Baruch, 2015) and in the United Arab Emirates (Ziegler, Debatin, et al., 2019).

As regards age and gender differences for the ten QELC subscales, a study with a large Turkish sample (1,620 students in groups of mean ages 10.5, 13.08 and 16.20 years) reached statistically significant sex differences only within the group of 13-year-olds, in favor of girls, for the economic, cultural, social, organismic and telic capitals. With respect to age, statistically significant differences appeared between the two other age groups (of 10-year-olds and 16-year-olds), with higher cultural capital scores for the younger students (Leana-Taşcilar, 2015).

The aims of the current study are two-fold, following a theoretical and also a methodological-statistical-metric perspective. The theoretical aims are the description of the QELC structure in a large Greek sample and its extension to exploring cross-cultural differences and similarities. The methodological and statistical aims are: i) the overall attempt to verify the theoretical QELC dimensions in this Greek sample via zero-order confirmatory factor analysis, ii) the implementation of specific statistical methods to test for modeling implications, iii) the application of a variation of an "exploratory SEM" attempt (Asparouhov & Muthén, 2009) to allow for comparisons across the first-order CFA outcomes and second-order confirmatory factor models, and iv) a cross-cultural comparison across correlation tables, in an initial attempt to compare the Greek QELC data with four other countries: China, Germany, Israel, and Turkey.

Method

Sample

Our sample consisted of 740 students, recruited from 16 Greek primary and secondary public schools in Athens and some other urban and non-urban areas of Greece. For the 50 QELC items, we first explored for possible missing values and 24 cases were detected in which only 80% or less of the items had been responded to (n of missing items ^{3} 10). Excluding these 14 cases from further analysis, all 726 remaining cases did not exceed six missing answers per case, all distributed randomly. These missing values were replaced by the corresponding variable mean, therefore for all 726 cases the full data set was available. To double-check this replacement for all 50 items, we compared the standard deviations before and

after replacement; the largest change reached only –.005 and its mean change was approx. –.002. Thus, the changes in standard deviations were trivial and we retained the replaced missing values for further analysis (n = 726).

In terms of the students' gender, location of their school and place of residence, 361 (49.7 %) were males and 365 (50.3 %) were females; 275 students (37.88 %) reside in Athens, 206 (28.37 %) reside in sub-urban areas and 245 (33.75 %) reside in rural areas. Two age groups were formed, the first one consisted of students of the fourth and fifth grades of primary school (ages 10 and 11, mean age = 10.46) and the second group included students of the sixth grade of primary school and also the first and second grades of junior high school (ages 12, 13 and 14, mean age = 12.56).

Instrumentation

The Questionnaire of Educational and Learning Capital (QELC), comprising 50 questions regarding students' educational capital (5 subscales) and learning capital (5 subscales) (Ziegler & Baker, 2013), was administered to students in classrooms. Each subscale (5 items) measures one of the ten forms of capital. The ten subscales are the following (an example-item for each subscale is given): *learning capital* includes the *organismic* subscale ("Being physically fit also helps me to learn and study for school for long periods of time"), the *actional* subscale ("I know a lot of strategies for learning and studying"), the *telic* subscale ("I set a goal for myself to continuously improve my performance at school"), the *episodic* subscale ("I have a lot of experience on how I can do very well in school"), and the *attentional* subscale ("In my daily routine, nothing distracts me from learning and studying for school"). *Educational capital* consists of the *economic* subscale ("My family spends more money on my schooling than other families do"), the *cultural* subscale ("I know a lot of people who think that learning and studying are very important"), the *social* subscale ("Other people give me good advice on how I can further improve my academic performance"), the *infrastructural* subscale ("Because of my good learning and studying environment I can be among the best in school"), and the *didactic* subscale ("During classroom instruction I am taught how to learn & study more effectively").

The QELC subscales show satisfactory to high Cronbach's α levels in the literature, ranging from .60 to .83, except for the telic subscale where α was only .49 (Vladut, Liu et al., 2013). For the sample of this study, Cronbach's α reliability estimates were satisfactory (> .73); specifically, these estimates by subscale were as follows: for the learning capital subscales, organismic subscale a reached .74, for the actional subscale α = .76, for the telic subscale α = .74, for the episodic subscale α = .81, and for the attentional subscale α reached .79; for the educational economic subscale a reached .80, for the cultural subscale α = .76, for the social subscale α = .77, for the infrastructural subscale α = .75, and for the educational didactic subscale α reached .85.

Results

Descriptive statistics, outliers, and the "appropriate coefficients" question

Basic descriptive statistics were first calculated at the item level. Although the measurement level of these 50 items is ordinal, we should briefly comment on the basic properties of these items; most distributions were negatively skewed (the average skewness was approx. –.82). However, for most of the items only less than 3 univariate outliers existed, and these outlier cases were more than ten for only one item, with the overall median value being 2 outlier cases. Thus, no action was taken with respect to outliers and all 726 cases were retained.

The second step was to examine whether *Pearson's r* correlations might possibly be affected by the skewness levels in our data; for this, we computed both parametric correlations (*Pearson's r*) among all 50 QELC items and also non-parametric ones (*Kendall's Tau-b*). Comparing these matrices through *Fisher's z* transformation (Mylonas, Veligekas, Gari, & Kontaxopoulou, 2012), only 90 out of the 1,225 non-parametric correlations were significantly different at the .05 level, from the corresponding parametric ones (approx. 7%). Thus, we were fully justified in employing parametric correlations in further analyses as the "appropriate coefficients hypothesis" was not refuted.

The age-group question; a developmental approach

The third step was of developmental nature and referred to the sample different developmental age bands (groups "1" to "5", corresponding to 10 to 14 years of age, respectively). This age variability begged the question of whether the initial raw scores might be suffering from bias due to age variations. If so, we should adjust the raw scores by extracting this bias from the measure (Mylonas & Furnham, 2014). This procedure, if deemed necessary, would affect the raw scores of the initially biased items themselves but not the correlations which would remain the same in any case. First a partial correlation approach was conducted and through *Fisher's z* transformation no statistically significant difference emerged between the partial correlations and the zero-order ones, despite the numerous arithmetic differences observed. Suspecting that true differences might be masked due to the nature of the measures, we also employed an eta-correlation approach and we computed all η and η^2 indices reflecting the correlation of two age-bands (10 & 11 vs. 12 to 14 years) with each of the 50 items. We indeed detected ten items with large η and η^2 values; for these, the method described by Mylonas & Furnham (2014) was employed and initial raw scores were adjusted by removing the unwanted variance (these items were named as 'cor' ones and appear as such in Tables and Figures).

Confirmatory Factor Analysis; 1st order modeling

The main 1st order CFA outcomes are summarized in Table 1.

After justifying the use of the available data through the procedures described above, we employed factor-analytic techniques to test for the Ziegler theoretical structure (Ziegler, Vialle, & Wimmer, 2013). We expected 25 items to form five separate quintuple factors on learning capital and the other 25 items to form five other separate quintuple factors on educational capital. We tested this model for the overall sample (n=726), as developmental effects were not present any more and there was no need to run the analysis separately for each age group. We computed the outcomes through LiSRel and R statistical packages. The CFA outcomes, as computed for successive models, are summarized in Table 1 and Figure 1.

The independence model was, as expected, not acceptable in terms of statistical fit. We then tested a unifactorial model with all 50-items being considered as manifesting one-single latent variable; this model was not accepted either, although some rather interesting properties were revealed, such as the nearly acceptable RMSEA, the acceptable SRMR, the elevated Tucker-Lewis Index (TLI) as compared to the null (independence) model, the large AIC and BIC reduction, and the large drop in the $\chi^2 \div df$ ratio along with the very large and statistically significant reduction in the χ^2 value itself. From these results, one might argue

Table 1 *First Order Confirmatory Factor Analysis: Summary of outcomes*

Model	χ^2 [i]	*df*	$\chi^2 \div df$	*RMSEA* [$CI_{90\%}$]	*SRMR*	*GFI*	*CFI*	*TLI*	$\Delta\chi^2$	*Δdf*	*AIC*	*BIC*
a	17,093.16	1,225	13.95	–	–	–	–	–	–		17,193.16	18,048.35
b	4,936.35	1,175	4.20	.080 [.078 - .082]	.060	.73	.76	a-b .76	a-b 12,229.81***	50	6,785.70	5,595.11
c	2,900.35	1,130	2.57	.050 [.048 - .052]	.042	.85	.89	b-c .50	b-c 1,963,00***	45	3,473.68	3,855.45
d	2,696.51	1,120	2.41	.047 [.045 - .050]	.041	.86	.90	c-d .10	c-d 203.84***	10	3,260.23	3,717.58
								b-d .55	b-d 2,166.84***	55		
e	4,742.30	1,130	4.20	.079 [.077 - .081]	.060	.74	.77	b-e ≈.00	b-e 121,05***	45	6,509.64	5,697.50

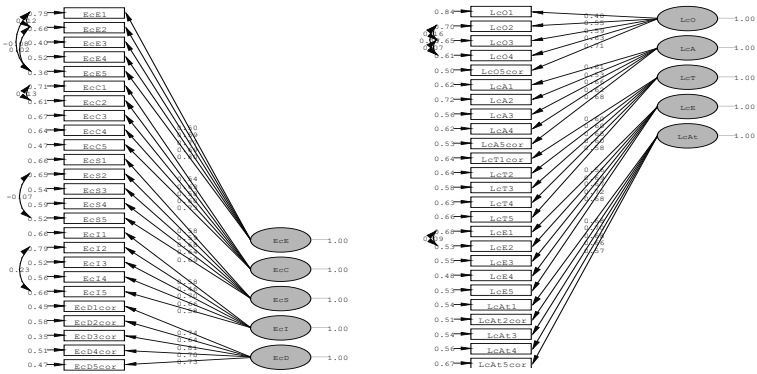
Key: a = Independence model, ***b*** = Unifactorial model, ***c*** = 10-factor model, ***d*** = modified 10-factor model, ***e*** = "random-structure" 10-factor model

[i] : All minimum fit function χ^2 criteria were significant at the .001 level

* significant at the .05 level, ** significant at the .01 level, *** significant at the .001 level.

Note A: 10 raw item-scores were corrected for developmental stage bias (see developmental approach under the 'age-group question' section) but correlations are the same and this correction has no effect on CFA outcomes.

Note B: *Error covariances* (model ***d***) strictly within factors: EcI5-EcI2, EcE2-EcE1, EcE5-EcE1, EcE2-EcE5, EcC2-EcC1, LcO4-LcO2, LcO3-LcO2, LcO3-LcO4, LcE2-LcE1, EcS5-EcS2



Key

EcE: Educational capital, ***Economic***
EcC: Educational capital, ***Cultural***
EcS: Educational capital, ***Social***
EcI: Educational capital, ***Infrastructural***
EcD: Educational capital, ***Didactic***

LcO: Learning capital, ***Organismic***
LcA: Learning capital, ***Actional***
LcT: Learning capital, ***Telic***
LcE: Learning capital, ***Episodic***
LcAt: Learning capital, ***Attentional***

Note: "cor" items have been adjusted for age-band bias

Figure 1 *Model "d ": CFA for 50 QELC Scores and 10 latent variables*

for a central, overarching theme present in the data, grouping all items under a single broader construct; we did not pursue this goal further in the current study, but future research may reveal interesting results.

The outcomes obviously suggested further modeling, so we then tested our main model (model "*c*"), that is the 10-factor theoretical structure for the five educational capital dimensions – economic, cultural, social infrastructural and didactic –and the five learning capital dimensions – organismic, actional, telic, episodic, and attentional ones (Ziegler, 2005; Ziegler, et al., 2013). Despite the notorious χ^2 statistical significance, the improvement with respect to the non-acceptable unifactorial model was evident in the AIC and BIC reduction, in the large TLI and in the statistically significant $\Delta\chi^2$ value. Although the fit for this model was good with CFI being nearly perfect, and RMSEA and SRMR reaching acceptable levels, the GFI reached only .85 and the $\chi^2 \div df$ value was still rather high, despite its large difference from the unifactorial model. As is obvious, some fine-tuning was necessary if we were to finally decide to retain this model, and these amendments should be carried out via the estimation of a few error covariances.We estimated 10 such error covariances, as were indicated during model "*c*" testing, allowing for error terms to correlate between specific items. We did so *strictly within factors*, thus not allowing for

any cross-loadings involvement in our modeling. For the modified model "*d*", a much better fit was observed, with CFI, RMSEA, and SRMR at acceptable levels, with TLI showing a large improvement with respect to the unifactorial model and acceptable improvement from the non-modified ten-factor one, with the respective $\Delta\chi^2$ values being statistically significant as well and with AIC and BIC being further reduced. The GFI value and especially the $\chi^2 \div df$ value were not at perfect levels though, however, this model ("*d*") had shown the best statistical fit so far and clearly supported the existence of the theoretical ten-factor structure in our data and was accepted as the best model at this stage. The loadings for this solution are presented in Table 2.

Still, some other questions remained; one of them was whether such improvement in statistical fit might be attributable to the large number of factors (10), especially considering the fact that we tested a unifactorial model and we then moved directly to the theoretical one which directs to numerous (ten) factors to be modeled. To account for such a possible method effect we also tested for a "random-structure" model, also with ten factors (model "*e*"). We randomly assigned items to the ten latent variables, modeling five random items for each factor (Table 1), so that each factor would enter the model as manifested from random combinations of learning and educational capital items (e.g., the telic dimension in the model was expected to be manifested from two organismic, one episodic, one didactic and one telic item). As expected, χ^2 was higher than the successful "*d*" model and the $\chi^2 \div df$ ratio returned to the unifactorial model levels. RMSEA was not acceptable anymore and CFI and GFI were very far from being acceptable.

Table 2 *Loadings; Solution = Model "d "*

Educational capital items	**Economic**	**Cultural**	**Social**	**Infrastructural**	**Didactic**
	(EcE)	(EcC)	(EcS)	(EcI)	(EcD)
1	.50	.54	.58	.58	.74
2	.59	.63	.59	.46	.64
3	.77	.58	.68	.70	.81
4	.69	.60	.64	.66	.70
5	.80	.73	.69	.58	.73
Learning capital items	**Organismic**	**Actional**	**Telic**	**Episodic**	**Attentional**
	(LcO)	(LcA)	(LcT)	(LcE)	(LcAt)
1	.40	.61	.60	.56	.68
2	.55	.53	.60	.69	.70
3	.59	.66	.65	.67	.68
4	.63	.62	.60	.72	.66
5	.70	.68	.58	.68	.57

A closer look though revealed that this "random-structure" model closely resembled our unifactorial modeling outcomes, in terms of goodness of fit indices and in terms of AIC, BIC, and χ^2 values. This seemed to underline the need to account for one or more overarching factors, a need which appeared but only in a subtle way during the unifactorial modeling. This, along with theoretical considerations, called for 2nd order factoring, as was carried out in the next stage.

Before following this quest though, we computed the aggregate scores for each of the ten dimensions to form new measures to be used in further analyses ("averaged aggregates"). These new measures reflected the five educational capital and the five learning capital factors; their basic descriptive statistics are presented in Table 3.

Table 3 *Descriptive Statistics for the Ten Aggregate Scores (Educational Capital and Learning Capital Factors)*

Educational Capital	***Economic***	***Cultural***	***Social***	***Infrastructural***	***Didactic***
	(EcE)	***(EcC)***	***(EcS)***	***(EcI)***	***(EcD)***
Mean	3.99	4.80	4.64	4.31	4.46
Median	4.00	5.00	4.80	4.40	4.70
Standard Deviation	1.20	.93	.94	.98	1.11
Standard error of Mean	.04	.03	.03	.04	.04
Skewness	−.49	−1.00	−0.85	−.46	−.80
Kurtosis	−.20	.96	.90	−.09	.17
Range	6.00	5.00	5.00	5.00	4.90
Min.	.00	1.00	1.00	1.00	1.07
Max.	6.00	6.00	6.00	6.00	5.97
Kolmogorov-Smirnov z *	.077	.124	.104	.081	.112
Shapiro-Wilk W *	.972	.926	.948	.977	.940

Learning Capital	***Organismic***	***Actional***	***Telic***	***Episodic***	***Attentional***
	(LcO)	***(LcA)***	***(LcT)***	***(LcE)***	***(LcAt)***
Mean	4.14	4.62	4.53	4.50	4.31
Median	4.20	4.80	4.60	4.60	4.40
Standard Deviation	1.00	.91	.97	.95	1.01
Standard error of Mean	.04	.03	.04	.04	.04
Skewness	−.41	−.97	−.84	−.80	−.58
Kurtosis	−.08	1.28	.79	.88	.09
Range	5.19	4.99	5.39	5.00	4.97
Min.	1.00	1.01	1.00	1.00	1.02
Max.	6.00	6.00	6.00	6.00	5.99
Kolmogorov-Smirnov z *	.073	.110	.105	.092	.077
Shapiro-Wilk W *	.982	.941	.951	.956	.969

* all statistically significant at the .001 level

Taking a closer look at the loadings for each of the items on its respective factor, these are not "1.00", obviously; one might wonder whether the computed aggregates truly reflect the dimensions as the weight for each participating item has been overlooked. Of course, we accept that their loading to other, theoretically unrelated factors is 0.00, and this was fixed in our modeling after all, but the relative importance of each item manifesting each factor is not the same. To explore for such a possible method effect, we computed the factor scores for this outcome (model "*d*") and we then correlated these scores with the aggregates (simple averaged sums). The results are presented in Table 4. One would expect only the diagonal elements to exceed .90 and under perfect similarity .95 (for averaged aggregates and factor scores to share at least 90% of their variance). However, this was not the case, and in three cases the correlation levels dropped even below .90. This irregularity was further highlighted by the fact that other, error correlations around .90 existed, between non-related factors, showing some possible levels of collinearity in the data, especially among the learning capital factors. This outcome underlined the need for 2nd order factor modeling, as is presented next.

Confirmatory Factor Analysis; 2nd order modeling

The first action was to note down as reference models the independence (null) model and the modified ten-factor model outcomes, as these were computed in the previous stage (1st order factoring). Models "*1*" and "*2*" are the same as the 1st order CFA "*a*" and "*b*", so their difference remains the same of course, but for model "*3*" (all models tested are summarized in Table 5) we introduced two second order factors to be tested, namely educational capital and learning capital, each comprising its respective five factors. We tested this solution further through an amended 2nd order factor model (model "*4*") also assuming orthogonality of all factors (model "*5*").

Table 4 *Correlations between Averaged Aggregates and Anderson-Rubin Factor Scores (fs)*

		fs_1	***fs_2***	***fs_3***	***fs_4***	***fs_5***	***fs_6***	***fs_7***	***fs_8***	***fs_9***	***fs_10***
EcE	1. Economic	***.975***	.501	.556	.505	.292	.464	.375	.414	.446	.380
EcC	2. Cultural	.491	***.951***	.813	.639	.512	.527	.640	.642	.620	.573
EcS	3. Social	.560	.843	***.930***	.802	.671	.637	.748	.743	.718	.676
EcI	4. Infrastructural	.499	.649	.786	*.897*	.769	.784	.785	.793	.758	.778
EcD	5. Didactic	.306	.549	.694	.816	***.969***	.687	.809	.821	.711	.808
LcO	6. Organismic	.443	.539	.621	.767	.626	***.933***	.751	.708	-.700	.746
LcA	7. Actional	.390	.690	.777	.834	.804	.820	*.882*	.910	.895	.865
LcT	8. Telic	.426	.681	.760	.830	.806	.762	.898	*.876*	.860	.888
LcE	9. Episodic	.457	.663	.736	.791	.698	.754	.881	.859	***.944***	.833
LcAt	10. Attentional	.388	.605	.691	.814	.789	.806	.852	.885	.832	***.932***

Table 5 *Comparing 1st and 2nd Order Factor Outcomes*

Model	χ^2 [i]	df	$\chi^2 \div df$	RMSEA [$CI_{90\%}$]	SRMR	GFI	CFI	TLI	$\Delta\chi^2$	Δdf	AIC [j1]	BIC [j2]
1	17,093.16	1,225	13.95	–	–	–	–	–	–		17,193.16	18,048.35
2	2,696.51	1,120	2.41	.047 [.045 - .050]	.041	.86	.90	[1-2] .89	[1-2] 14,396.65***	105	3,260.23	3,717.58
3	3,295.37	1,172	2.81	.050 [.048 - .052]	.064	.83	.87	[1-3] .86	[1-3] 13,797.79 ***	53	3,501.37	3,973.89
4	3,200.37	1,159	2.76	.049 [.047 - .051]	.048	.84	.87	[1-4] .86	[1-4] 13,892.79 ***	66	3,432.37	3,964.53
5	4,144.96	1,173	3.53	.059 [.057 - .061]	.219	.82	.82	–	–	–	12,304.96	31,022.16

Key: 1 = Independence model, ***2*** = modified 10-factor model (1st order CFA model "*d*") , ***3*** = second order factor model (two second order factors), ***4*** = amended 3rd model, see note A below, ***5*** = orthogonal factors model, see Note B below

[i] : All minimum fit function χ^2 criteria were significant at the .001 level

[j1] : Computed as $\chi^2 + \{ [k(k + 1)] - 2df \}$

[j2] : Computed as $\chi^2 + \{ \ln(N) \{ [k(k + 1) \div 2] - df \} \}$

* significant at the .05 level, ** significant at the .01 level, *** significant at the .001 level.

Note A: The *4th* model is the *3rd* model amended for negative parameter estimates

Note B: This orthogonal model assumes zero correlation between latent variables (1st and 2nd order ones)

The 3rd model is obviously not as good as the 2nd one which is the 1st order modified ten-factor model. Although the RMSEA and the SRMR indices remain at acceptable levels, the GFI and CFI values have dropped and the $\chi^2 \div df$ ratio is enlarged. The gain (TLI) from the independence model is less, as compared to model "*2*" and the χ^2 itself is larger. Finally, AIC and BIC values are also larger. Thus, a direct 2nd order factor structure does not seem to fit the data well. We also detected negative parameter estimates during the computations, a multicollinearity side-effect. To remedy this, and in an attempt to enhance the solution, we attempted to amend this model by relaxing the initial parameters for the 2nd order factors and for the reference factors in the first-order solution. The outcome was slightly better (model "*4*"), but still the $\chi^2 \div df$ ratio was larger than the one in model "*2*" and χ^2 itself was larger as well; however, the GFI, CFI, RMSEA and SRMR indices remained at acceptable levels, with a tendency to match the indices in model "*2*" (1st order model "*d*"). Finally, a 5th model was tested with orthogonal latent variables. The RMSEA index remained at acceptable levels, but the GFI and the CFI values were much lower than acceptable levels; the $\chi^2 \div df$ ratio was enlarged along with the χ^2 value itself. The most problematic indices in this solution were the SRMR (.219), along with the large AIC an BIC values; thus this orthogonal model did not fit the data whatsoever, an indication that there is some inter-connection both among the 1st order factors and between the 2nd order ones, with the same indication present in the 1st order solutions as depicted in the correlational analysis of factor scores and averaged aggregates showing some extent of a dynamic "osmosis-like" connection between educational and learning capitals. Finally, for the amended 2nd order factor solution and for reasons of comparability with the 1st order

factoring solution, we present (Table 6) the loadings which were computed for model "*4*" in this stage of analysis. Finally, model "*4*" loadings of the 1st order factors on the 2nd order ones can be found in Figure 2 which graphically shows this model's outcomes.

Table 6 *Loadings; Solution = Model "4 "*

Educational Capital items	***Economic***	***Cultural***	***Social***	***Infrastructural***	***Didactic***
1	.49	.59	.58	.60	.74
2	.62	.66	.58	.48	.64
3	.77	.57	.69	.70	.81
4	.69	.59	.64	.67	.70
5	.78	.72	.67	.60	.73
Learning Capital items	***Organismic***	***Actional***	***Telic***	***Episodic***	***Attentional***
1	.42	.61	.60	.58	.69
2	.63	.54	.61	.69	.70
3	.63	.66	.65	.68	.68
4	.69	.61	.60	.73	.66
5	.70.	.68	.59	.68	.56

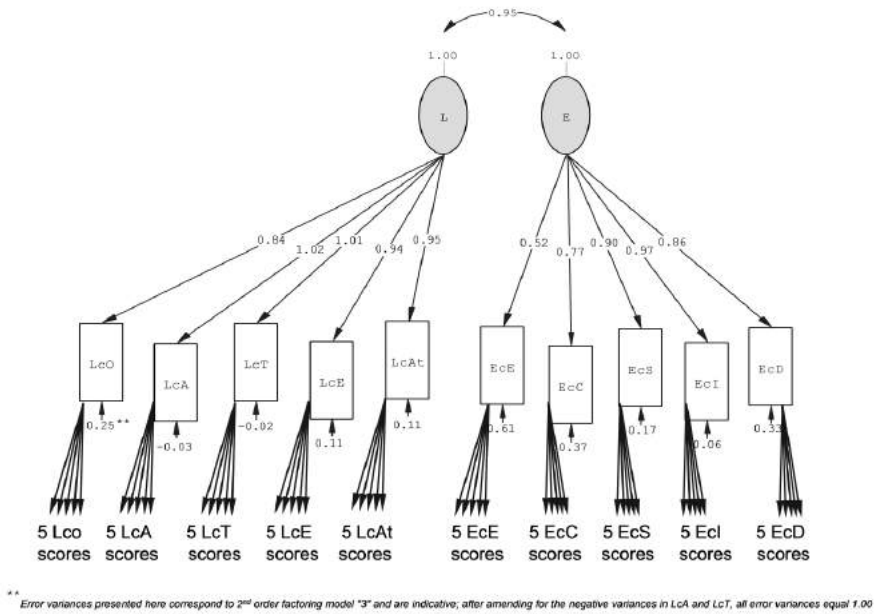


Figure 2 *Model "4": 2nd Order CFA for 2 Higher-Order Factors, 10 First Order Ones, and 50 QELC scores*

A cross-cultural approach; correlation similarity across four countries

For the last step of our analysis, we employed four zero-order correlation tables for the 50 QELC-items as published for China, Germany, Israel and Turkey (Paz-Baruch, 2015; Vladut, Liu et al., 2013). Having computed the same correlation table for Greece, we compared all five correlation tables through *Fisher's z* transformation in order to explore for possible differences in the patterns of item inter-correlations. The existence of such differences, if present, might indicate important "guidelines" for future cross-cultural modeling (i.e., through multi-group confirmatory factor analysis and/or multivariate covariance structure analysis). The outcomes are summarized in Table 7. The numbers below the diagonal refer to the absolute number of significantly different pairs of correlations ($p < .05$) between countries and the entries above the diagonal are the respective percentages.

We should first note that for the 10 pairs of inter-correlation matrices (correlating the averaged aggregate QELC scores) almost all of them were different to a larger or a smaller extent (statistically significant differences exceeding 15 % of the correlation pairs); the only exception was observed between the Israeli and German matrices which seem at least quite similar. The Chinese correlational pattern (cp) differs from those of all the other four countries (31 % to 49 % of correlation pairs are different), and especially from the German and the Turkish ones. Israeli cp is marginally different from the Greek one (16 %) but clearly different from the Turkish one (22 %). Greek cp is quite different from both the German and the Turkish one (29 %), while the Turkish cp is quite different from the German one (22 %), as well.

Age and place differences

A series of one-way ANOVAs were conducted, in order to explore the relations of some demographic and other variables of interest with the 10 QELC capitals. Specifically, we related these scores with gender, two age groups (9- & 10-year-olds vs. 12- to 14-year-olds), and students' permanent place of residence and school location (capital vs. provinces). Comparing across places, the educational didactic capital mean was higher for Athens ($M = 4.81$) compared with $M = 4.15$ for provincial places ($F_{1, 724} = 70.63$, $p < .001$ [$p = .000$], $\eta^2 = .09$), but the learning telic capital mean score was higher for students of provincial areas ($M = 4.73$), as compared with Athens ($M = 4.35$), $F_{1, 724} = 28.03$, $p < .001$ [$p = .000$], $\eta^2 = .04$.

Comparing across age bands, the didactic capital mean was higher ($M = 4.79$) for the

Table 7 *Correlation Comparisons (Through Fisher's z, 45 Coefficients) Across Five Countries*

	China	***Germany***	***Greece***	***Israel***	***Turkey***
China	0	49 %	33 %	31 %	44 %
Germany	22	0	29 %	13 %	22 %
Greece	15	13	0	16 %	29 %
Israel	14	6	7	0	22 %
Turkey	20	10	13	10	0

younger age group, in comparison with the older age group (M = 4.25) ($F_{1, 724}$=43.93, p < .001 [p = .000], η^2 =.06). The same held true for the attentional capital; mean score was higher (M= 4.58) for the younger ages, than the older (M = 4.13) ($F_{1, 724}$=36.63, p < .001 [p = .000], [], η^2 =.05). No differences were found in terms of gender.

Discussion

In line with previous research results for the QELC reliability and the relevant theoretical assumptions, we verified the ten factor CFA model for a Greek sample, including five distinct factors for educational capital and five factors for learning capital (Ziegler & Baker, 2013; Ziegler, Balestrini, et al., 2018; Ziegler, Debatin, et al., 2019; Ziegler & Stoeger 2017). All ten forms of learning capital – organismic, actional, telic, episodic, and attentional, along with the educational economic, cultural, social, infrastructural and didactic capitals, seem to co-exist, as distinct internal and external types of resources. The interdependent co-existence of these capitals seem to form a dynamic potential for the school life span of all students; their effective combination seems a necessary and also a sufficient prerequisite in order to fulfill any high potential demand, regardless of the specific starting point for each student. Within a functional and flexible motivating environment, an interaction of these ten capitals can nurture an excellent system of outputs (Ziegler & Phillipson, 2012) not only for gifted students but for the great majority of students. Nevertheless, educational and learning capitals, as core issues within the Actiotope Model of Giftedness, seem to re-orient the study of giftedness towards a systemic understanding of excellence. It may also re-define the "key starting point" to support *all students* at school, towards an amelioration of their potential into the highest possible abilities and skills.

Our results seem to corroborate the existing evidence that the QELC is indeed a reliable and metrically valid tool for cultural use and possibly cross-cultural comparisons, under -emic and -etic perspectives (Hui & Triandis, 1985), enriching our cross-cultural view of the giftedness concept and the identification of gifted students. We should not fail to comment on some levels of (multi) collinearity present in our data though. This was observed mainly within the learning capital subscales, especially while we correlated the averaged aggregate scores with the factor scores directly computed from the factor structure's factor coefficients and through the osmosis indications during 2nd order CFA modeling. Although this collinearity does not seem to pose a threat to the scale's validity, we need to reflect on this osmosis that possibly works under the surface and inter-relates several learning capital subscales. The phenomenon does not appear within the educational capital subscales, possibly due to stronger independence powers reflecting a more clear-cut distinction of the constructs in the students' minds; this distinction may not be so clear with respect to learning capital constructs though, possibly reflecting the many vague, unlimited or undifferentiated aspects pervading the Greek educational system, which can result in further undifferentiated conceptions of it by the students (Ziegler, Debatin, et al., 2019; Ziegler & Phillipson 2012; Ziegler & Stoeger, 2017).

Cultural and cross-cultural studies on giftedness, which are still scarce (Stoeger, Balestrini, et al., 2018), can be empowered

via further cross-cultural QELC studies. The current study's initial cross-cultural results reflect the relations among the ten Greek educational and learning capitals as compared to the correlational patterns among the ten capital relations in China, Israel, Germany, and Turkey (Leana-Taşcilar, 2015; Paz-Baruck, 2015; Vladut, Liu, et al., 2013). The greatest similarities appeared between students in Israel and Germany, but the greatest differences appeared between students in China and students in all the other four countries. For the remaining students (Israel, Greece, and Turkey), all correlational patterns of relations differed amongst them and all other countries, either to a larger extent, e.g., the Greek patterns with the German and Turkish patterns or to a smaller extent e.g. the Turkish patterns with the Israeli and German ones, or the Israeli patterns with the Greek ones. In general, in this comparison of four different cultural settings, it is important to note that only one "cultural pattern" of correlations was common across Germany and Israel. Obviously, these findings cannot be fully interpreted at this stage of analysis, but they may offer some important hints for further research. In this future attempt, and ideally under a large – or at least larger – number of participating cultures, we might attempt a multivariate exploration of the similarities and differences across these correlational patterns relating educational and learning capitals through methods such as the "hit-matrix" and MDS-T methods (Georgas & Mylonas, 2006; Mylonas, 2009; Papazoglou & Mylonas, 2017), as described and applied in other studies on family values. Through these methods we might be able to better understand the common grounds and cross-cultural differences and/or even form homogeneous clusters of cultures sharing strong similarities in the way students capitalize on education and learning potential.

Finally, and further extending the above, cultural variety of educational and learning capitals across groups *within* a country e.g., among social groups of cultural and linguistic diversity, may be an obvious research necessity, as has been supported in a methodological study (Mylonas, 2009). In the current study, an initial effort to depict cross-cultural differences between groups within the Greek social setting was conducted through our comparisons between the Athenian subsample and the provincial areas one. Such an attempt may possibly indicate some of the variables to be considered as intra-country sub-groupings, possibly allowing or even demanding a "cross-cultural" analytic lens in future research. Specifically, the differences found for the educational didactic capital favored students in Athens, but for the learning telic capital, the differences found favored students in provincial towns; this finding apart from being an interesting point in our current discussion, also seems to be an interesting "hint" for further research and interpretation, under the within country "cross-cultural" rationale, as "place of residence" can culturally differentiate across groups of students (Gari, Mrvoljak, & Nikolopoulou, 2019 *April*). An initial interpretation may associate the higher didactic capital for students in Athens, (a capital city with a population of 3,762,000), with the higher information access opportunities these students enjoy, along with more alternative educational material, teaching methods and teaching "experts". On the other hand, the higher telic capital for students in smaller cities/towns and provincial places, as compared with students in the biggest city of the

country, may depict their higher need to set life goals of specific meaning, and to pursue them in order to satisfy their present and future needs; otherwise, they may remain restricted within the limited provincial chances; these limited chances for "educational goods and commodities" may greatly urge them to self-regulate learning and to set for themselves alternative goals towards more effective learning and life chances (Ziegler, Stoeger, & Grassinger, 2011).

Limitations and future research

A first warning should stem from our attempt to reduce age-band inflicted bias, as this was triggered by the developmental method-factor, possibly active in any QELC data. The adjustment itself may be a discussion point, as one might wonder if we should continue adjusting the scores for other sources of bias as well; however, this might either lead us to unwantedly reduce the variance to levels less than adequate for statistical analysis, or in contrast would simply reflect some haphazard selection of exogenous factors to neutralize, which in the end would have devastating effects on the observed factorial structure, an effect we obviously wanted to avoid. Thus, one has to carefully select the biasing factors to control for, but apart from that, we should also note the actual property in our data, which forced us to adjust the scores, and this is no other than the developmental aspect itself. In this study and through the adjustments we performed, we avoided modeling our data separately for different age-groups, but this may not always be feasible or even desirable. The developmental factor should always be considered in future research so as to be properly treated as a part of the model or as a possible confound.

Some sample irregularities were related to the above concerns. One of them was the fact that age differences amongst participants existed – an advantageous disadvantage really. Another concern regards place of residence; although for the purposes of the current study the respective distribution can be considered satisfactory, if one wanted to draw safer conclusions about the necessity of including place of residence, either as a method factor or as a correlate in any multivariate modeling, one would have to better represent urban, sub-urban, and rural populations in Greece, also by testing for the possible departures from such a population distribution through the appropriate χ^2 tests. Of course, for the current non-normative study, the above do not constitute a serious threat, but in future research, they definitely should not be neglected. In addition, age-bands are not a perfect way to explore for possible correlations between educational-learning capital items and age; it would be much better to have the students' birth dates available so that their precise age (in days) might be computed and used as a correlate; this would definitely add to our power while removing the unwanted variance in the QELC data possibly caused by age variations. It would also be an idea for future studies either to keep the age constant in the sample, or the opposite, to study larger age spans, to either diminish any possible biasing effects or to study them in greater depth, respectively.

In future research, further comparative studies may also be conducted, inspired by prior eco-cultural and eco-social modeling; these models showed extensive functionality in depicting cultural similarities and differences across countries, or across cultural groups within a country (Georgas & Berry,

2005; Georgas, van de Vijver, & Berry, 2004); utilizing a set of a priori selected specific eco-social indices that are associated with five educational and five learning capitals dimensions, we might end up with a circumplex of differentiated clusters of countries/cultures of obvious theoretical and applied value.

References

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling, 16*, 397–438. https://doi.org/10.1080/10705510903008204

Bateson, G. (1979). Mind and Nature: A Necessary Unity. Dutton. https://monoskop.org/images/c/c3/Bateson_Gregory_Mind_and_Nature.pdf

Csikszentmihalyi, M. (1996). *Creativity: Flow and the psychology of discovery and invention*. Harper Collins Publishers.

Csikszentmihalyi, M. (1988). Society, culture, and person: A systems view of creativity. In R. J. Sternberg (Ed.), *The nature of creativity* (pp. 325-339). Cambridge University Press.

Dowling, E. (1985). Theoretical framework – a joint systems approach to educational problems with children. In E. Dowling & E. Osborne (Eds.), *The family and the school: A joint systems approach to problems with children* (pp. 5-32). Routledge and Kegan Paul. https://doi.org/10.1002/1520-6807(198907)26:3<318::AID-PITS2310260315>3.0.CO;2-3

European Commission (2016). Education and Training Monitor 2016. *Greece Education and Training*. https://ec.europa.eu/education/sites/education/files/monitor2016-el_en.pdf

Gari, A. (2016). The dynamics of gifted students' psychoeducational provisions in Greece. *ECHA News, 30*(1), 11-12. https://www.echa.info/images/documents/news/ECHA%20News%20vol%2030%20no%201%202016.pdf

Gari, A. (2019). An in depth view of what happens to the gifted and talented students in Greece. *33*(1), 14-15. https://www.echa.info/images/documents/news/ECHA_News_vol_33_no_1_2019.pdf

Gari, A., Kalantzi–Azizi, A., & Mylonas, K. (2000). Adaptation and motivation of Greek gifted pupils: exploring some influences of primary schooling. *High Ability Studies, 11*:1, 55-68.

Gari, A., Mrvoljak, I., & Nikolopoulou, V. (2019, *April 12-16*). *Greek student's strengths, difficulties and self-efficacy correlated with educational and learning capital: Teachers' and Students' Perceptions* [Paper presentation]. 6th Biennial Conference of the International Research Association for Talent Development and Excellence (IRATDE), Taipei, Taiwan.

Georgas, J. & Berry, J. (1995). An ecocultural taxonomy for cross-cultural psychology. *Cross-Cultural Research, The Journal of Comparative Social Science, 29*(2), 121-157. https://journals.sagepub.com/doi/10.1177/106939719502900202

Georgas, J. & Mylonas, K. (2006). Cultures are like all other cultures, like some other cultures, like no other culture. In U. Kim, K.S. Yang & K.K. Hwang (Eds.), *Indigenous and Cultural Psychology: Understanding People in Context* (pp. 197-221). New York: Springer.

Georgas, J., van de Vijver, F., & Berry, J. W. (2004). The ecocultural framework, ecosocial indices, and psychological variables in cross-cultural research. *Journal of Cross-Cultural Psychology, 35*(1) 74-96. https://journals.sagepub.com/doi/10.1177/0022022103260459

Hui, C. H. & Triandis, H. (1985). Measurement in cross-cultural psychology: A review and comparison of strategies. *Journal of Cross-Cultural Psychology, 16*, 131-152. https://doi.org/10.1177%2F0022002185016002001

Leana-Taşcilar, M. Z. (2015). Age differences in the Actiotope Model of Giftedness in a Turkish sample. *Psychological Test and Assessment Modeling, 57*(1), 111-125. https://www.academia.edu/11683743/Age_differences_in_the_Actiotope_Model_of_Giftedness_in_a_Turkish_sample

Molnar, A. (1986). A Systemic Perspective on Solving Problems in the School. *NASSP Bulletin, 70*(493), 32-40. https://doi.org/10.1177/019263658607049307

Mönks, F. J., Heller, K. A., & Passow, A. H. (2000). The study of giftedness: reflections on where we are and where we are going. In Heller, K. A., Mönks, F. J., Sternberg, R. J., & Subotnik, R. F. (Eds.), International Handbook of Giftedness and Talent (2nd Ed.) (pp. 839-863). Pergamon.

Mönks, F. J., Van Boxtel, H. W., Roelefs, J. J. W. & Sanders, M. P. M. (1986). The identification of gifted children in secondary education and a description of their situation in Holland. In K. A. Heller & J. F. Feldhusen (Eds.), *Identifying and nurturing the gifted* (pp. 39-65). Hans Huber Publishers. https://epub.ub.uni-muenchen.de/2422/1/2422.pdf

Mylonas, K. (2009). Statistical analysis techniques based on Cross-Cultural research methods: cross-cultural paradigms and intra-country comparisons. *"Psychology": The Journal of the Hellenic Psychological Society, 16*(2), 185-204.

Mylonas, K. & Furnham, A. (2014). Bias in Terms of Culture and a Method for Reducing It: An Eight-Country "Explanations of Unemployment Scale" Study. *Educational and Psychological Measurement, 74*(1), 77-96. https://10.1177/0013164413502669

Mylonas, K., Veligekas, P. Gari, A., & Kontaxopoulou, D. (2012). Development and Psychometric Properties of the Scale for Self-Consciousness Assessment. *Psychological Reports: Measures and Statistics, 111*(1), 233-252. https://doi.org/10.2466%2F08.02.07.PR0.111.4.233-252

Papazoglou, S., & Mylonas, K. (2017). An examination of Alternative Multidimensional Scaling Techniques. *Educational and Psychological Measurement, 77*(3), 429–448. https://doi.org/10.1177/0013164416661823

Paz-Baruch, N. (2015).Validation study of the Questionnaire of Educational and Learning Capital (QELC) in Israel. *Psychological Test and Assessment Modeling, 57*(2), 222-235. http://www.psychologie-aktuell.com/fileadmin/download/ptam/2-2015_20150624/06_Paz-Baruch.pdf

Phillipson, N. S., S., Phillipson, & Eyre, D. (2011). Being gifted in Hong Kong: An examination of the region's policy for gifted education. *Gifted Child Quarterly 55*(4), 235-249. https://doi.org/10.1177/0016986211421959

Stoeger, H., Balestrini, D. P. & Ziegler, A. (2018). International perspective and trends in research on giftedness and talent development. In S. I. Pfeiffer (Ed.), *APA Handbook of Giftedness and Talent* (pp. 24-37). *American Psychological Association* https://psycnet.apa.org/doi/10.1037/0000038-002

Tannenbaum, A. J. (1983). *Gifted children: Psychological and educational perspectives*. Macmillan.

Tannenbaum, A. J. (1988). Probing Giftedness/ Talent/ Creativity: Promise and Fulfillment. In P. F. Brandwein & A. H. Passow (Eds.), *Gifted Young in Science. Potential through Performance,* (pp. 39-55). https://eric.ed.gov/?id=ED306128

Tannenbaum, A. J. (2000). A History of Giftedness in School and Society. In K. A. Heller, F. J. Mönks, R. F. Sternberg, & R. F. Subotnik, *International Handbook of Giftedness and Talent,* 2nd edition (pp. 23-53). Elsevier.

Vladut, A., Vialle, W., & Ziegler, A. (2015). Learning resources within the actiotope: A validation study of the QELC (Questionnaire of Educational and Learning Capital). *Psychological Test and Assessment Modeling, 57*(1), 40-56. https://www.psychologie-aktuell.com/fileadmin/download/ptam/1-2015_20150327/03_Vladut.pdf

Vladut, A., Liu, Q., Leana-Taɑcilar, M. Z.,Vialle, W., & Ziegler, A. (2013). A cross-cultural validation study of the Questionnaire of Educational and Learning Capital (QELC) in China, Germany and Turkey. *Psychological Test and Assessment Modeling, 55*(4), 462-478. https://www.psychologie-aktuell.com/fileadmin/download/ptam/4-2013_20131217/08_Vladut.pdf

Von Bertalanffy, L. (1968). *General System Theory: Foundations, Development, Applications*, New York: George Braziller. https://monoskop.org/images/7/77/Von_Bertalanffy_Ludwig_General_System_Theory_1968.pdf

Ziegler, A. (2005). The Actiotope Model of Giftedness. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (pp. 411-434). Cambridge University Press. https://doi.org/10.1017/CBO9780511610455

Ziegler, A. & Baker, J. (2013). Talent development as adaptation: The role of educational and learning capital. In S. Phillipson, H. Stoeger, & A. Ziegler (Eds.), *Exceptionality in East Asia: Explorations in the Actiotope Model of Giftedness* (pp. 18-39). Routledge.

Ziegler, A., Balestrini, & Stoeger, H. (2018). An international view on gifted education: Incorporating the micro-systemic perspective. In S. I. Pfeiffer (ed.), *Handbook of giftedness in children,* Springer International Publishing. https://doi.org/10.1007/978-3-319-77004-8

Ziegler, A., Chandler, K., Vialle, W., & Stoeger, H. (2017). Exogenous and endogenous learning resources in the Actiotope Model of Giftedness and its significance for gifted education. *Journal for the Education of the Gifted, 40*, 310-333. https://doi.org/10.1177/0162353217734376

Ziegler, A., Debatin, T. & Stoeger, H. (2019). Learning resources and talent development from a systemic point of view. *Annals of the New York Academy of Sciences, 1-13,* https://doi.org/10.1111/nyas.14018

Ziegler, A. & Phillipson, S. (2012). Towards a systemic theory of giftedness. *High Ability Studies, 23*, 3-30. https://doi.org/10.1080/13598139.2012.679085

Ziegler, A. & Stoeger, H. (2008). A learning oriented subjective action space as an indicator of giftedness. *Psychology Science Quarterly*, *50*, 222-236.
https://www.researchgate.net/publication/26519065_A_learning_oriented_subjective_action_space_as_an_indicator_of_giftedness

Ziegler, A. & Stoeger, H. (2017). Systemic gifted education. A theoretical introduction. *Gifted Child Quart*erly, *61*, 183–193. https://doi.org/10.1177%2F0016986217705713

Ziegler, A., Stoeger, H., & Grassinger, R. (2011). Actiotope model and self-regulated learning. *Psychological Test and Assessment Modeling, 55*(1), 161-179. https://www.psychologie-aktuell.com/fileadmin/download/ptam/1-2011_20110328/08_Ziegler.pdf

Ziegler, A., Stoeger, H., Harder, B., Park, K., Portešová, S., & Porath, M. (2014). Gender differences in mathematics and science: The role of the actiotope in determining individuals' achievements and confidence in their own abilities. *High Ability Studies, 25*, 35-51. https://doi.org/10.1080/13598139.2014.916092

Ziegler, A., Vialle, W. & Wimmer, B. (2013). The actiotope model of giftedness: A short introduction to some central theoretical assumptions. In S. N. Phillipson, H. Stoeger & A. Ziegler (Eds.), *Exceptionality in East Asia* (pp. 1-17). Routledge. https://ro.uow.edu.au/sspapers/526

Corresponding author:
Aikaterini D. Gari, PhD
Department of Psychology
National and Kapodistrian University of Athens
E-mail: agari@psych.uoa.gr

Validation of the Educational and Learning Capital Questionnaire (QELC) on the Mexican population

Grecia Emilia Ortiz Coronel[a], María de los Dolores Valadez Sierra[a], María Elena Rivera Heredia[b], María de Lourdes Vargas Garduño[b], Óscar Ulises Reynoso González[a], Jaime Fuentes Balderrama[c]

[a] University of Guadalajara
[b] Michoacana University of San Nicolas of Hidalgo
[c] University of Texas

Abstract:
The Actiotope Model of giftedness is a systemic model with a focus on actions directed towards objectives of ability development. As such, the development of talents and extraordinary achievements is considered an intelligent adaptation to environmental and personal stimuli. The Questionnaire of Educational and Learning Capital (QELC) may allow for empirical evidence of successful adaptation of an actiotope. Thus, the purpose of this study was to examine the psychometric properties of The Questionnaire of Educational and Learning Capital in the Mexican population. A total of 374 gifted Mexican elementary school students participated ($\bar{X}$=11.18 age, S.D. 1.36). We calculated its internal consistency and performed confirmatory factor analysis. The results show that the original factor structure presents absolute fit, and low levels of error. Additionally, we observed adequate values of extracted variance (0.5<AVE) and composite reliability (0.7 < CRI) on all factors except for the economic subscale.

Keywords:
high capacities, student body, QELC, validity, CFA

Validation of the Educational and learning Capital Questionnaire (QELC) on the Mexican population

The Mexican education system is considered one of the largest in the world. It is composed of preschool, elementary school and middle school. In its three levels, preschool education focuses on children three to five years old, elementary education incorporates children six to twelve years old and consists of six grades, and middle school education teaches three grades to young men and women 13 to 15 years of age. Of the total student body that attends basic education levels (elementary and middle school), it is estimated that ten percent of the students have high capacities. Despite the fact that attention towards these students has been proposed, it has not been carried out in all of the schools in the country and many students frequently go unnoticed by the usual identification processes (Secretary of Public Education [SEP], 2006, 2017, 2019).

Although there is an intervention proposal, Educational Attention for Students with Outstanding Capacities by the Secretary of Public Education of Mexico and there is a clearly stated processes for identification and intervention, it is important to have other instruments that allow for the consideration of the personal and social factors associated with an individual. After all, it is very important that education for the students with the greatest capacities reaches the entire student body.

Students that are identified as possessing high capacities have academic, social and emotional experiences that are related to their individual resources and their environment. This generates great challenges in different dimensions (García-Barrera & de la Flor, 2016). The problems that the twenty-first century society needs to solve regarding people with high capacities and talents go beyond the intellectual coefficient or cognitive capacity. A way to approach how we should identify these individuals is by asking what challenges the world encounters at a given point in time. In other words, to really understand the performance of people with high capacities, it is necessary to understand the resources that can be used according to the demands presented by their surroundings (Covarrubias, 2018; Sternberg, 2017).

Ritchotte (2013) states that for decades researchers have studied high capacities in an effort too help gifted students reach their maximum capacity and prevent potentially devastating consequences such as school desertion. Thus, it is important to have models that explain and identify individuals with high capacities that allows for the comprehension of their resources and favors efficient attention for this population.

Nowadays there is a diversity of concepts and explicative models (performance-based models, sociocultural orientation models, cognitive models, and capability-based models). Each one studies a series of specific characteristics, which makes identification confusing and ambiguous. None of these explicative models of high capacities is able to encompass, with all of its interactions, a definition, study method and an education proposal that corresponds to all of the realities of society (Ziegler, Vialle, & Wimmer, 2013).

In the last decade different researchers such as Renzulli, Gagné, Tannenbaum, Mönks and Gardner have contributed to the understanding of gifted individuals through the expansion of both the individual and social fields of action. This is the case of the

actiotope model of giftedness, which aids in the identification of the resources that are partially found in the student (endogenous resources) and outside of the student (exogenous resources). This is a systemic model, which means that all of its elements interact (Ziegler & Stoeger, 2017). Actions are directed towards the development of abilities. As such the development of talent and extraordinary achievements is considered an intelligent adaptation to environmental stimuli (Vladut, Vialle, & Ziegler, 2015; Ziegler et al., 2013).

The actiotope model of giftedness focuses on the actions of an individual and its evolution. The development of excellence is understood as a dynamic system adaptation that intensifies in complexity through the interactions with the objective structure of a dominion. Thus, with increasing excellence, the individual will also achieve greater changes in the objective structure of self-dominion. The model takes into account the co-adaptation and co-evolution of the components of the Actiotope such as the range of actions and determinants, goals, subjective space of action and environment and the interpretation of these components within a network. These are gifts and talents that are traditionally understood as attributes of an individual (Ziegler, 2005).

Ziegler and colleagues (Ziegler & Baker, 2013; Ziegler, Chandler, Vialle, & Stoeger, 2017; Ziegler, Debatin, & Stoeger, 2019) suggest that the regulation of endogenous resources is subjected exclusively to the subsystem of the "person". However, even though exogenous resources may be used by the person, its supply generally depends on other systems (i.e. school, teachers, piers, educational system). They associate exogenous resources with the term educational capital, and the term endogenous resources with learning capital.

To offer empirical evidence for the model, we administered a low-cost instrument that measures two general factors: Educational Capital and Learning. Each factor contains five subscales (i.e. educational capital: economic, cultural, social, infrastructure and didactic; learning capital: organic, actional, telic, episodic and attentional). This instrument is called The Questionnaire of Educational and Learning Capital (QELC) developed by Vladut, Liu, Leana-Taşcilar, Vialle, and Ziegler, (2013) and adapted by Leana-Taşcılar (2016) for teachers. First studies demonstrated satisfactory psychometric qualities across different cultures, demonstrating content and construct validity (Vladut et al., 2013; Leana-Taşcılar, 2016).

Furthermore, Vladut, et al. (2015) adapted the QELC for elementary and middle school students. The results show that the reliability of the ten QELC subscales had a satisfactory range. However, reliability was low for the subscale that measures actional learning capital (α = 0.62) and telic learning capital (α = 0.68). The CFA model fits to the data when the five forms of educational capacity are influenced by a latent variable and the remaining five forms of learning capacity are influenced by a second latent variable.

In Israel, Paz-Baruch (2015) evaluated the validity of the QELC with a sample of 187 elementary school students from Israel to examine if the educational and learning capital of the students was related to intelligence and academic achievement. The study found correlations between infrastructure, didactic, organic, actional, episodic and attentional capacity. No correlation was found between general intelli-

gence and other QELC subscales. The internal consistency results and bifactorial CFA model confirmed the validity and reliability of learning and educational capital using the Hebrew version of the QELC.

Based on the validity and confidence results obtained by the above-mentioned studies, this research has the objective of understanding the psychometric properties of The Questionnaire of Educational and Learning Capacities to validate its use on the Mexican population.

Method

Participants

A convenience sampling was chosen, due to the availability for the application of the instrument, this ex post facto study recruited Mexican students from seven primary schools in the city of Guadalajara (n = 374), of which 102 students belong to the public sector and 272 students to the private sector. The age of the participants was $\bar{X}$ = 11.18 (S.D. 1.36); 50.5% of the participants (n = 189) were women and 49.5% were men (n = 189) (table 1).

Instrument

The Questionnaire of Educational and Learning Capital (QELC) (Vladut et al., 2013) is a 50 item-long self-report (e.g. I know from experience how to learn better), which is answered with a 6-point Likert-type scale, ranging from 1 = "Strongly disagree" to 6 = "Strongly agree". The 50 items are grouped into 10 subscales divided into two factors: Educational (Economic [1, 11, 21, 31, 41], Cultural [2, 12, 22, 32, 42], Social [3, 13, 23, 33, 43], Infrastructure [4, 14, 24, 34, 44], and Didactic [5, 15, 25, 35, 45]) and Learning (Organic [6, 16, 26, 36, 46], Actional [7, 17, 27, 37, 47], Telic [8, 18, 28, 38, 48], Episodic [9, 19, 29, 39, 49] and Attentional [10, 20, 30, 40, 50]). The QELC has demonstrated construct validity as well as acceptable internal consistencies (α = .57 <.86) in Germany, Turkey, Israel and China (Paz-Baruch, 2015; Vladut et al., 2013), yet the QELC's psychometric properties have not yet been examined in Spanish-speaking countries.

Table 1. *Characteristics of the participants*

Characteristics	*n*	%
Sex		
Female	185	49.5%
Male	189	50.5%
Grade		
4th grade	50	13.4%
5th grade	106	28.3%
6th grade	87	23.3%
1st middle school	97	25.9%
2nd middle school	16	4.3%
3th middle school	18	4.8%

Analysis

To identify whether the QELC is a valid instrument for Mexican samples, a Confirmatory Factor Analysis (CFA) was carried out.

To adapt the QELC from German to Spanish, we used the guidelines proposed by Beaton et al., (2000) for the translation and semantic equivalence of the items. The adaptation is composed by four phases: translation to Spanish, correspondence analysis, cultural adaptation and empirical adaptation.

Prior to the CFA, frequency distributions of the items and subscales were reviewed to assess their normality. Similarly, internal consistency analyses were carried out using the subscales reported by the authors. Complementary to analyzing the factorial solution of Vladut et al., (2013), we present the analysis when constraining to a single factor instead of two, as well as using the 10 subscales as first-order factors and the Educational and Learning scales as second-order factors.

Given the absence of negative multivariate kurtosis and approximate multivariate normality, the CFA were carried out using the default Maximum Likelihood estimator and fit was assessed using the following indices: For absolute fit we used the χ^2 statistic, where a non-significant discrepancy value is expected; for close fit we used CFI, TLI and GFI, where values above 0.9 indicate good fit and values above 0.95 indicate excellent fit (Abad et al., 2011; Hair et al., 1999). To evaluate acceptable error values, we used the SRMR and the RMSEA; SRMR values < 0.10 are considered acceptable, whereas for RMSEA values < 0.08 indicate an acceptable fit, and values < 0.05 indicate very good fit. To quantify fit discrepancy between models, we used the AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) as comparative criteria, where smaller values indicate better fit (see table 2) (Browner & Crudeck, 1993; Byrne, 2001; Littlewood & Bernal, 2011; Moral, 2006). Additionally, the Average Variance Extracted (AVE) and the Composite Reliability Index (CRI) were calculated. Values of AVE > 0.5 indicate an adequate percentage of explained variance and, values of CRI > 0.7 are sufficient (Hair et al., & Black, 1999). Preliminary analyses were carried out in SPSS V.25 and the CFA was performed in AMOS V.24.

RESULTS

Analysis of normality and internal consistency

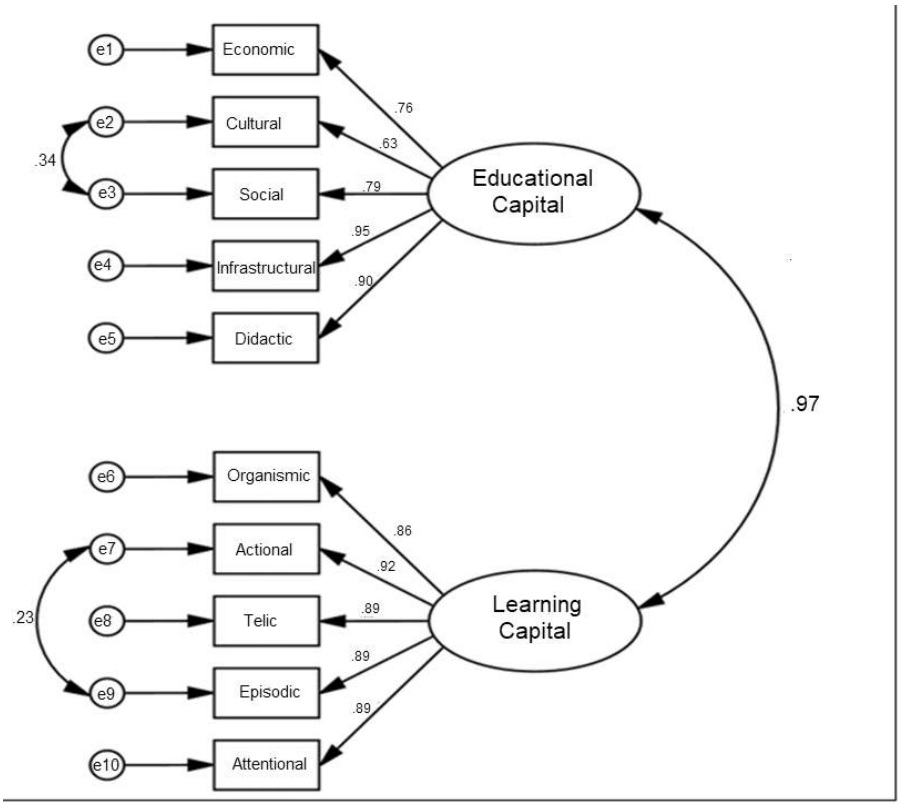
Although only seven items presented non-significant Shapiro-Wilk values, none of the items or subscales presented asymmetry or kurtosis values greater than the cut-off points (|2| and |6| respectively), which suggests approximately normal distributions. With the exception of "Economic", the subscales showed moderate to high internal consistencies (α=0.79 < 0.86) and the consistencies of the factors were excellent: Educational (α = 0.94) and Learning (α = 0.96). Item 31 (i.e. "I think my education is very expensive") had a low total correlation with the rest of the items in the "Economic" subscale (r = 0.21), which threatened internal consistency. By eliminating this item, the internal consistency reached acceptable levels (α = 0.59 to α = 0.64). Subsequent analyses were run including and excluding item 31 as a sensitivity analysis, yet fit indices, determination coefficients and internal consistency always favored the exclusion of this item from the analyses.

Confirmatory Factor Analysis

The CFA was carried out using the factorial solution reported by Vladut et al., (2013), where five subscales correspond to each of the two factors (Educational and Learning). We complementarily ran an analysis constraining all subscales to one factor, and another model with second order factors where each subscale is estimated as first order latent factors.

The original model by Vladut et al., (2013) demonstrates close fit as well as moderate levels of error (see Table 2: Model 1). Modification indices do not suggest cross-loadings, all factorial loadings are significant, latent variables present significant variances and the variances explained for the subscales are high ($R^2 = 0.41 < 0.89$) (see figure 1). Given the high correlation between both factors, another analysis was run linking the 10 subscales to a single factor but the chi-squared fit index demonstrated a sig-

Figure 1. *Questionnaire of Educational and Learning Capital (QELC) (Vladut et al., 2013). Model 4: Original factorial solution with two Modification Indices.*



Source: Original work based on AMOS output

nificantly worse fit when compared to the first model $\Delta\chi$ (1) = 31.62 p <.000, ruling out this alternative (see Table 2: Model 2).

To improve the fit of the original model, the analysis was rerun adding the Lagrange multiplier with the largest coefficient (covariance between the errors of the Cultural and Social subscales of the Educational factor) and subsequently with the modification index presenting the greatest magnitude for the other factor (covariance between the errors of the Episodic and Actional subscales, now from the Learning factor). Both additions significantly improved absolute fit ($\Delta\chi(1) > 3.84$ in both cases), increased close fit and decreased residual errors with respect to the previous model (Higher CFI, TLI, and GFI, lower SRMR, RMSEA as well as AIC and BIC). Our fourth model presented excellent close fit, as well as acceptable levels of error (see Table 2: Model 4). Although it does not present absolute fit, the discrepancy is probably due to the absence of strictly normal data. Despite our approximately normal univariate distributions, small fluctuations in bias and kurtosis would violate the assumption of multivariate normality posed by the Maximum Likelihood estimator, which represents an obstacle in approaching absolute fit (Curran et al., 1996).

Complementarily, each subscale was independently analyzed as a first order factor to later be combined and thus validate the instrument with first and second order factors. To avoid an overfitting of the sample, only two Lagrange multipliers were allowed per subscale. The fit indices for these analyses are presented in Table 3.

All first-order factors presented absolute fit and low levels of error. Additionally, adequate values of variance extracted (0.5 <AVE) as well as composite reliabilities (0.7 <CRI) were observed in all the factors, except the Economic factor (even without item 31). When estimating the second-order factor "Educational", only the Cultural, Social and Didactic sub-factors were formed. In contrast, Economic and Infrastructure are presented as Heywood cases (Byrne, 2001) (i.e. Negative variances and factor loadings greater than one). In a base model with only the 5 first-order factors, the correlations between factors were unexpectedly high (r =

Table 2. *Fit indices from the Confirmatory Factor Analyses*

Model	χ^2	*gl*	*p*	CFI	TLI	GFI	SRMR	RMSEA	AIC	BIC
Model 1: QELC	138.5	34	.0001	.97	.96	.92	.02	.091*** [.075- .11]	180.6	263.09
Model 2: QELC single factor	170.2	35	.0001	.96	.95	.91	.03	.102*** [.09- .12]	210.7	289.27
Model 3: Model 1 with one M.I.	95.3	33	.0001	.98	.97	.95	.02	.071* [.055- .088]	140.7	249.05
Model 4: Model 1 with two M.I.	83.5	32	.0001	.98	.98	.96	.02	.066 [.049- .083]	131.5	221.83

Note: *M.I.= Modification Index.*

Table 3 *Fit indices for first-order factors*

Factor	χ^2	*gl*	*p*	*CFI*	*TLI*	*GFI*	*SRMR*	*RMSEA*	*AVE*	*CRI*
Educational Capital										
Economic	2.01	2	.365	1.00	1.00	.99	.01	.004 [.000- .103]	.37	.67
Cultural	5.30	3	.151	.99	.98	.99	.01	.045 [.000- .107]	.51	.77
Social	6.45	3	.091	.99	.98	.99	.01	.056 [.000- .115]	.65	.84
Infraestructure	5.46	5	.362	.99	.99	.99	.01	.016 [.000- .075]	.68	.85
Didactic	1.56	5	.906	1.00	1.00	.99	.00	.000 [.000- .029]	.68	.85
Learning Capital										
Organic	5.50	5	.357	.99	.99	.99	.01	.000 [.000- .029]	.67	.85
Actional	1.72	3	.632	1.00	1.00	.99	.00	.000 [.000- .070]	.62	.83
Telic	5.36	3	.147	.99	.98	.99	.01	.046 [.000- .108]	.61	.82
Episodic	2.85	4	.582	1.00	1.00	.99	.01	.000 [.000- .067]	.68	.85
Atttentional	4.40	3	.221	.99	.99	.99	.01	.000 [.000- .029]	.56	.80

Note: AVE=Average Variance

0.65 < 0.97 vs. r = 0.55<.83 when they were sumscales) and even aberrant for the pairs Infrastructure with Didactics (r = 1,001) and Economic with Infrastructure (r = 1,006), which suggests multicollinearity and a poor specification of the model where each of the 10 indicators are treated as a first-order factor. The second order factor "Learning" presented the same problem since the Actional and Attentional factors were also presented as Heywood cases. In the base model, correlations between factors were still higher than for the previous factor (r = 0.86<.99 vs. r = 0.73 < 0.86 when they were sumscales), and the correlation between the Actional and Episodic pair was aberrant (r=1.003), also suggesting multicollinearity and a poor specification of the model if the subscales are taken as first order factors. In this sense, the original factorial solution reported by Vladut et al. (2013) presented construct validity in this Mexican sample, as well as moderate and high levels of internal consistency.

Discussion

Nowadays, nations strive to provide some of the development standards to produce scientists in different fields. Thus, these countries attempt to discover gifted children and present them with adequate attention to help them to be creative in the future (Olimat, 2010)

In Mexico, the identification of children with high capacities and the attention given to them has had important developments. However, the actions employed in this country and other parts of the world are

still insufficient. Efforts are isolated, and previously generated resources and experiences are underutilized (Valadez, 2019). In Mexico the identification and evaluation of these students is frequently expensive due to parents, professors and specialists implementing different resources. Identification is through exploratory activities by teachers and through products of the children's aptitudes. Meanwhile evaluation is a series of standardized psycho-pedagogical tests that include the application of intelligence tests, creativity and socialization.

As was mentioned earlier, there are various models of explanation and giftedness and talent. However, none of them are capable of encompassing, with all of its interactions, a definition, a study method and an educational proposal, that matches all of the realities of society (Ziegler et al., 2013).

However, the Actiotope Model of Giftedness is a systemic model with a focus on directed actions towards objectives concerning the development of abilities. As such, the development of talents and extraordinary achievements is considered an intelligent adaptation to environmental stimuli (Vladut et al., 2015).

Instead of identifying individuals through classic cognitive methods (tests of intelligence coefficients), it analyzes the route of entry that learning and excellence will have (Leana-Taşcılar, 2016). These resources are partially found in the student (endogenous resources) and partly found outside of the student (exogenous resources).

Ziegler et al. (2017) suggest that the regulation of endogenous resources is subject exclusively to the "person" and that the person can use exogenous resources. Their provision generally depends on other systems. Thus, they associated the term of learning capital with endogenous resources. Based on this model, Vladut et al. (2013) developed a questionnaire that is a quantitative economic measurement instrument that allows for large-scale surveys on students. The resulting instrument, the QELC, includes only 50 items grouped into 10 subscales (five for educational capital and five for learning capital), and was designed as transculturally applicable product (Leana-Taşcılar, 2016).

The objective of this study was to understand the psychometric properties of the QELC instrument to validate the theoretical assumptions of the Actiotope Model of Giftedness in the Mexican population. However, this study was not without limitations. The sample was collected from seven schools in Guadalajara through non-probabilistic methods meaning our results should be interpreted in light of these conditions and its generalizations should be done with caution.

Once the items were adapted from the original scale through the direct translation method, we modified the items in the economic subscale from the education capital with the help of the authors of the questionnaire. The changes were requested because the perception of insecurity in Mexico regarding violence is a public problem that makes the people's quality of life more vulnerable. In Mexico, more than half (66.1 %) of the populations feels insecure in the state in where they live. This has stopped people from doing everyday activities, which has repercussions on social recreation and relaxation, and inhibits social cohesion, even generating fear when providing information on personal economy (Jasso, 2013).

Afterwards we revised the psychometric properties of the instrument through confirmatory factor analysis. We found evidence for satisfactory psychometric prop-

erties, which confirms the construct validity of the QELCS's original factorial solution proposed by Vladut et al., (2013) in Mexico. However, we observed difficulties regarding the Economic factor, particularly with the adjustment of item 31.

We tested three distinct factorial solutions: 1) the original model proposed by the authors, 2) A single latent factor model, and 3) an exploratory model devised from using the 10 sumscales as first-order latent factors which belong to two second-order factors. Our results support the original model with two covarying first-order factors and ten sumscales based on 49 items (not the original 50, as we eliminated item 31 due to reliability purposes). By eliminating item 31 (i.e., "I think that my education is very expensive") we obtained better psychometric propertied. This is because the subfactor "Economic" of educational capital is strongly related to exogenous resources and closely linked to the violent conditions of Mexico such as kidnapping, theft and organized crime. However, all in all the QELC demonstrates satisfactory psychometric qualities in the Mexican context and provides a valid representation of the existing relationships between the variables that make up the educational and learning capitals for high-capacity elementary school students in Guadalajara, Mexico.

References

Abad, F., Olea, J., Ponsoda, V., & García, C. (2011). *Measurement in social and health sciences.* Madrid: Synthesis.

Beaton D., Bombardier C., Guillemin F., & Ferraz, M. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)*, *25*(24), 186-91. doi: 10.1097/00007632-200012150-00014

Browner, M. W, & Crudeck, R. (1993). Alternative ways of assessing model fit. *Testing structural equation models, 21*(2), 230-258. doi:10.1177/0049124192021002005

Byrne, B. M. (2001). *Structural equation modeling with AMOS: Basic concepts, applications and programming.* Mahwah, NJ: Erlbaum.

Covarrubias, P. (2018). From the concept of outstanding aptitudes to that of high abilities and talent. *Journal of Educational Research of the REDIECH, 9*(1), 53-67.

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods, 1*, 16–29.

Garcia-Barrera, A., & de La Flor, P. (2016). Spanish teachers' perception of students with high abilities. *Pedagogical Studies, 42*(2), 129-149. doi: 10.4067/S0718-07052016000200008.

Hair, J., Anderson, R., Tatham, R., & Black, W. (1999). *Multivariate analysis.* Madrid: Prentice Hall.

Jasso, C. (2013) Perception of insecurity in Mexico. *Magazine of Mexican Perception of Opinion.* 15, 12-29. DOI: 10.1016/S1870-7300(13)72319-6

Leana-Taşcılar, Z. (2016). Questionnaire of educational and learning capital (QELC): Turkish language validity and Factor Structure. *College Student Journal*. 531-543. Retrieved from https://toad.halileksi.net/sites/default/files/pdf/questionnaire-o-f-educational-and-learning-capital-qelc-toad.pdf

Moral, J. (2006). *Confirmatory factor analysis (CFA)*. México: Trillas.

Olimat, M. (2010). Gifted students. *Procedia - Social and Behavioral Sciences,* 5, 1112-1114. doi.org/10.1016/j.sbspro.2010.07.244.

Ritchotte, J. A. (2013). The validity of the achievement-orientation model for gifted middle school students. *ProQuest One Academic, 58(3),183-198.* doi:10.1177/0016986214534890

Secretary of Public Education. (2006). *Intervention proposal: Educational attention to students with outstanding with outstanding aptitudes*. Retrieved from http//: normalista.ilce.edu.mx

Secretary of Public Education. (2017). *Educational model. Equity and inclusion.* Retrieved from http://www.sems.gob.mx/work/models/sems/Resource/12302

Secretary of Public Education. (2018). *Main figures of the national education system.* Retrieved from *https://www.planeacion.sep.gob.mx/Doc/estadistica_e_indicadores/principales_cifras/principales_cifras_2017_2018_bolsillo.pdf*

Secretary of Public Education. (2019). *Main figures of the national education system.* Retrieved from https://www.planeacion.sep.gob.mx/Doc/estadistica_e_indicadores/principales_cifras/principales_cifras_2018_2019_bolsillo.pdf.

Sternberg, R. (2017). ACCEL: A new model for identifying the gifted. *Roeper Review, 39*(3), 152-169. doi: 10.1080 / 02783193.2017.1318658.

Valadez, M. D. (2019). Educational response to students with high abilities in Mexico. In C. B. Zarza, J. Delgado, R. De Araújo, & E. Morais (Eds.), *Brazilian and Mexican perspectives on the care of students with high abilities: linking theory and practice* (pp. 112-134). New York: Marlis Book.

Vladut, A., Liu, Q., Leana-Taɑcilar, M. Z.,Vialle, W., & Ziegler, A. (2013). A cross-cultural validation study of the Questionnaire of Educational and Learning Capital (QELC) in China, Germany and Turkey. *Psychological Test and Assessment Modeling, 55(4)*, 462-478. https://www.psychologie-aktuell.com/fileadmin/download/ptam/4-2013_20131217/08_Vladut.pdf

Vladut, A., Vialle, W., & Ziegler, A. (2015). Learning resources within the Actiotope: A validation study of the QELC (Questionnaire of Educational and Learning Capital). *Psychological Test and Assessment Modeling, 57*(1), 40–56.

Ziegler, A. (2005). The Actiotope Model of Giftedness. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (pp. 411-434). Cambridge University Press. https://doi.org/10.1017/CBO9780511610455

Ziegler, A. & Baker, J. (2013). Talent development as adaptation: The role of educational and learning capital. In S. Phillipson, H. Stoeger, & A. Ziegler (Eds.), *Exceptionality in East Asia: Explorations in the Actiotope Model of Giftedness* (pp. 18-39). Routledge.

Ziegler, A., Chandler, K., Vialle, W., & Stoeger, H. (2017). Exogenous and endogenous learning resources in the Actiotope Model of Giftedness and its significance for gifted education. *Journal for the Education of the Gifted, 40*, 310-333. https://doi.org/10.1177/0162353217734376

Ziegler, A., Debatin, T. & Stoeger, H. (2019). Learning resources and talent development from a systemic point of view. *Annals of the New York Academy of Sciences*, 1-13, https://doi.org/10.1111/nyas.14018

Ziegler, A. & Stoeger, H. (2017). Systemic gifted education. A theoretical introduction. *Gifted Child Quarterly, 61*, 183–193. https://doi.org/10.1177%2F0016986217705713

Ziegler, A., Vialle, W. & Wimmer, B. (2013). The actiotope model of giftedness: A short introduction to some central theoretical assumptions. In S. N. Phillipson, H. Stoeger & A. Ziegler (Eds.), *Exceptionality in East Asia* (pp. 1-17). Routledge.

Author Note

Grecia Emilia Ortiz Coronel, Interinstitutional PhD in Psychology, University of Guadalajara; María de los Dolores Valadez Sierra, Applied Psychology Department, University of Guadalajara; María Elena Rivera Heredia & María de Lourdes Vargas Garduño, Researcher-Professor at Michoacana University of San Nicolas of Hidalgo; Óscar Ulises Reynoso González, Research professor at the University of Guadalajara; Jaime Fuentes Balderrama, Research Associate Steve Hicks School of Social Work University of Texas at Austin.

The results of this research are part of the doctoral thesis of Grecia Emilia Ortiz Coronel, carried out with the support of the Mexican National Council of Science and Technology (CONACyT).We have no conflicts of interest to disclose.

Corresponding author:
María de los Dolores Valadez Sierra, PhD
University of Guadalajara
Sierra Mojada 950
Guadalajara, México
E-mail: dolores.valadez@academicos.udg.mx

Moderating Effects of Educational and Learning Capital on the Consequences of Performance Feedback

Anamaria Mendl[a], Bettina Harder[a], Wilma Vialle[b]

[a] University of Erlangen-Nuremberg, Germany
[b] University of Wollongong, Australia

Abstract:
Feedback is highly recommended in educational settings and can deliver valuable information to guide one's learning efforts. However, it can—but does not need to—lead to detrimental effects on learners' motivation and emotions. The response should be moderated by the individual's attitudes, experiences, abilities and many other aspects which can be subsumed in the concept of educational and learning capital. This study investigated the effects of random feedback on a reasoning test, measuring the participants' subsequent cognitive, metacognitive, motivational and emotional responses. Initially, we assessed their educational and learning capital to analyze direct and moderating effects on the responses. Results showed that feedback effects were not as strong as expected while educational and learning capital exerted strong direct effects. Moderating effects were verified for only some responses.

Keywords:
Educational and learning capital, feedback, performance, motivation, emotion

Moderating Effects of Educational and Learning Capital on the Consequences of Performance Feedback

Feedback is omnipresent in our educational system as well as in our workplaces; it can be summative or formative, mandatory or optional, delivered in various forms and at different frequencies. In the educational context, it often serves as a teaching tool that is anticipated to have an effect on the future performance of the learner. However, the effects can extend beyond changing performance, for example by influencing the learner's motivation, expectations, or emotional state (e.g., Bangert-Drowns, Kulik, Kulik, & Morgan, 1991; Kluger & DeNisi, 1996; O' Leary & O' Leary, 1977). The current study sought to determine the effects of feedback in a university learning context and how these effects could be moderated by the students' resources, namely their educational and learning capital.

Effects of Feedback

The body of evidence on the effects of feedback is inconsistent. On the one hand, feedback has been shown to improve learning performance (e.g., Jacobs, 2002; Krause & Stark, 2004) and, if given in an informative manner, it can increase the learner's intrinsic motivation (e.g., Deci, Koestner & Ryan, 2001; Sansone, 1986). Indeed, Hattie's (2009) meta-analysis of learning success in the school system concluded that feedback was one of the most important instruments to support learning processes (10^{th} of 100+). On the other hand, no effects or even negative effects of feedback on the learning process can be found (e.g., Jacobs, 2002; Kluger & DeNisi, 1996). For example, Kluger and DeNisi (1996) found that feedback that diverted the learner's attention to their self-concept—such as praise, verbal feedback and feedback that could threaten the self-concept—can lead to negative effects because the learner's cognitive resources no longer concentrated solely on the task, but also on themselves. Further, strong emotional reaction can hinder performance. In contrast to this, feedback that focuses on the task level should have more positive effects on performance.

According to learning theories, feedback can have cognitive (e.g., performance standards), metacognitive (e.g., validation of self-evaluation) and motivational effects (e.g., higher motivation; cf. Krause, 2007; London, 2002, 2003). These three effect groups (cognitive, metacognitive and motivational) are highly connected with each other in theory, so that an empirical differentiation may be challenging (Krause, 2007). Constructivist approaches also consider feedback effects on emotions (e.g., anxiety; cf. Kluger & DeNisi, 1996).

We will outline the different effects of feedback on (a) cognition, (b) metacognition, (c) motivation and (d) emotion in the following.

Feedback Effects on Cognition

Feedback has a diagnostic function, it helps one to clarify misunderstandings and to identify gaps in knowledge and skills, and it also indicates one's strengths and can emphasize the possible areas of improvement (Krause, 2007). In addition, if feedback is perceived as useful, it has a positive effect on achievement (Harks, Rakoczy, Hattie, Besser, & Klieme, 2014). For example, the study of Balzer, Doherty and O'Con-

nor (1989) shows that cognitive feedback improves the performance on judgment tasks. Further, the study of Vollmeyer and Rheinberg (2005) shows, that feedback improves cognitive performance: the feedback group had gained more knowledge than the no-feedback group.

Feedback Effects on Metacognition

Studies on feedback effects on metacognition have often considered metacognitive judgments (e.g., Geurten & Meulemans, 2016), self-regulated learning (e.g., Butler & Winne, 1995), and metacognitive skills (e.g., Molin, Haelermans, Cabus, & Groot, 2020). Consequently, feedback can validate judgements. Evaluations from others can stimulate the deliberate reflection of one's self (self-judgements) and of one's errors (error-judgments). In the context of teaching, error-judgements are particularly interesting, as coping with failure is relevant for self-regulated learning.

Coping with failure means not giving up after failure, but instead trying to regain control of the situation and to initiate new actions to attain the desired goal (Mantzicopoulos, 1990; Mietzel, 2005; Newton & Keenan, 1985). Dresel, Schober, Ziegler, Grassinger and Steuer (2013) differentiate two types in coping with failure: the affective-motivational adaptivity of error reactions (e.g., preservation of positive affects) and the action adaptivity of error reactions (e.g., adjusting learning activities in order to cope with the error). For both these error reactions, feedback is needed in order to handle errors and to be able to learn from errors.

Feedback Effects on Motivation

Based on previous findings, Vollmeyer and Rheinberg (2005) recommend teachers to use feedback, not only because it provides the learner with information, but also because feedback is beneficial for one's motivation. Many studies have investigated effects of (negative) feedback on intrinsic motivation (e.g., Fong, Patall, Vasquez, & Stautberg, 2018). However, effects of feedback on self-confidence and implicit personality theories are also of interest in the motivational context.

The self-concept of one's abilities comprises the assumptions of the level of one's abilities, which depend significantly on the attribution of learning outcomes (e.g., Weiner, 1986). Attribution can be steered by feedback, helping students to find the reasons for their positive or negative learning outcome. It is favorable for motivation to attribute successful learning outcomes to high abilities and personal effort, and in turn attribute failure to little effort (Dresel & Ziegler, 2006). Further, the attribution of failure to one's insufficient abilities is non-beneficial for motivation in general and the self-concept in particular (Dresel & Ziegler, 2006). The latter also plays a significant role for successful learning outcomes, and therefore many training concepts focused on fostering a positive self-concept, however, the outcomes were limited because beliefs about the modifiability of one's abilities were neglected by such trainings (Dresel & Ziegler, 2006).

Therefore, in addition to considering self-concept, it is important to assess and to provide feedback according to the different implicit personality beliefs, especially the beliefs about the modifiability of one's abilities that individuals might have (e.g.,

Dweck, 1999). According to this, individuals perceive their abilities and intelligence either as stable (entity view) or recognize them as adaptable (incremental view). For example, Mueller and Dweck (1998) found out that praising children for being smart would probably favor the development of a fixed mindset (entity view), which implies that these children would be less persistent after failure and their learning outcomes would worsen. In contrast, if children were praised for their hard work, this would promote the development of a growth mindset (incremental view), which entails the belief that they are able to improve after failure, thus leading to more persistence and better learning outcomes.

In contrast to Dweck's theory (e.g., 1999, 2006), Ziegler and Stoeger (2010) considered it necessary to refine the distinction between entity view and incremental view. They assume that the entity view only shows negative consequences if a person shows ability deficits, but has positive consequences if a person shows high abilities. Furthermore, Ziegler and Stoeger (2010) give an overview on the literature on self-theories pointing out the benefits of maintaining the stability of positively judged aspects of one's self. For instance, higher levels of self-confidence are not only associated with higher achievement, but also with better learning processes, and more adaptive behavior (e.g., Dweck, 1999; Gist & Mitchell, 1992; Pajares, 1996). Following this line of thought, Ziegler, Fidelman, Reutlinger, Vialle, and Stoeger (2010) proposed two beliefs about one's own learning: believing in the stability of one's abilities and believing in the modifiability of one's ability deficits. A stability belief refers to the conviction that one's learning reliably leads to success across a wide variety of situations and, thus, the result of one's learning is stable. Modifiability beliefs regarding one's ability deficits allow for improvement through learning and therefore make future success possible.

Feedback Effects on Emotion

Emotions also play an important role in the context of learning processes and their outcomes. Performance-related emotions should be considered in teaching (Pekrun, Elliot, & Maier, 2006) and therefore also in feedback situations. Yet, there has been little research on this topic (Värlander, 2008). Tennant (1997) draws attention to the fact, that once emotions are aroused, they cannot be turned off that easily, and they might hinder the learning process for days.

A perfect example is anxiety, possibly one of the strongest basic human emotions. According to the Yerkes-Dodson-law, performance increases with physiological or mental arousal, but only up to a certain point (e.g., Yerkes & Dodson, 1908). When anxious feelings overly dominate, especially in achievement situations, they can be dysfunctional towards learning outcomes. Kluger and DeNisi (1996) define anxiety as the "evaluation of a threat to the goals of the self, combined with a tendency to act to terminate the threat" (p. 267). Taking action to terminate the threat requires resources. Hence, feedback-induced emotion may influence the way the individual's available resources are used (Kluger & DeNisi, 1996).

According to self-enhancement theories, when faced with self-esteem relevant information, such as the case of feedback on one's performance, individuals tend to protect or increase their self-esteem (Peterson, Stahlberg, & Dauenheimer, 2000; Sedikides & Strube, 1995). When facing negative feedback, individuals tend to avoid

or reinterpret this information, in order to protect their self-esteem and to avoid negative emotions, such as anxiety. In contrast to self-enhancement theories, self-consistency theories (e.g., Swann, 2011) have a different explanatory approach whereby individuals tend to preserve the consistency of their self-attitudes. This implies that individuals with a positive self-esteem will prefer positive feedback, and that those with a negative self-esteem will prefer negative feedback. Empirical evidence for both approaches can be found, the effects depend on characteristics of the feedback (feedback design) and on personal characteristics of the learner (see section on moderating factors on feedback effects).

Our study focused on concepts from all four groups of feedback effects: (a) cognitive, that is, cognitive performance after feedback; (b) metacognitive, that is, coping with failure in two ways, namely adapting affective-motivationally to the failure and adapting one's actions after failure; (c) motivational, that is, self-confidence and stability/modifiability beliefs after feedback; and (d) emotional outcomes, that is, anxiety. This comprehensive overview of the feedback effects is the basis to determine possible moderating effects of resources which will be outlined next.

Moderating Factors on Feedback Effects

In the following sections, we will describe why the different effects of feedback on the learner and the learning outcomes can be explained by feedback design and through personal characteristics of the feedback recipient.

Feedback Design

Feedback design contains many different aspects, including form, content, timing, accuracy, reference norm and presentation mode of the feedback.

Feedback Forms

Some literature on feedback differentiates several forms of feedback (e.g., Jacobs, 2002; Kulhavy & Stock, 1989). Arguably, the most important differentiations are: knowledge of results (information about the achieved results), knowledge of correct results (information about the correctness of results), answering until correct (the learner has to try giving the answer until reaching the correct one), and elaborated feedback (additional information, for example, why the answer is correct or false). Elaborated feedback is more effective for subsequent performance than the other forms of feedback, particularly regarding complex tasks (Bangert-Drowns et al., 1991; Collins, Carnine & Gersten, 1987; Krause & Stark, 2004).

Feedback Content

Feedback information can refer to one's performance in a task or refer to oneself, such as personal development, or a combination of both. According to Kluger and DeNisi (1996), the effectiveness of feedback decreases as the feedback information becomes more self-orientated and through this, the attention of the feedback recipient moves away from task content towards themselves.

Feedback Timing

A further distinction can be made based on the timing of feedback. The terms "immediate" and "delayed" feedback are not precisely defined (Dempsey, Driscoll, & Swindell, 1993). Immediate feedback, for example, can be given after each item or after finishing the test, while delayed feedback can be given after a fixed time interval, such as a few seconds after finishing the test or after one week. According to the meta-analyses of Kulik and Kulik (1988), immediate feedback is on average slightly more effective for learning outcomes than delayed feedback. This positive effect of immediate feedback is stronger for simple tasks such as learning new words in a foreign language. When it comes to more complex tasks such as comprehending the content of a text, however, delayed feedback is more effective.

Feedback Accuracy

Feedback can be genuine, meaning the actual performance is fed back, or it can be fake, meaning a false performance is fed back. In cases of random feedback, which might be given to avoid performance levels confounding with feedback effects, special attention should be paid to accuracy so that the feedback can either match or mismatch the actual performance. Therefore, additional control variables are in order, such as assessing participants' awareness of the (in-)accuracy (e.g., Shibata, Yamagishi, Ishii, & Kawato, 2009) or asking for their performance judgements to assess their feedback expectation (e.g., Alloy & Abramson, 1979; Sedikides, Campbell, Reeder, & Elliot, 1998).

Feedback Reference Norm

When giving feedback, one can choose among (or also combine) the three different reference norms: criterial, individual and social reference norm (Krause, 2007). Many feedback studies use the criterial reference norm, where the content of feedback is orientated at the learning objective. When choosing the individual reference norm, feedback is orientated at the individual development in a learning domain. Using the social reference norm, students receive feedback on how good their individual performance was compared to other students. Findings show that, when using the criterial or the individual reference norm, feedback can be accepted more easily, than when using the social reference norm (Kopp & Mandl, 2014).

Feedback Presentation Mode

Feedback can be received face-to-face from others, such as teachers, other students, or researchers, or it can be received by non-person means, such as paper-pencil-based or computer-based feedback. Further, feedback can be presented in written, graphic or audible form. Findings suggest that, face-to-face feedback is easier to accept, because there is the possibility to request more information (Krause, 2007).

Personal Characteristics of the Feedback Recipient

As mentioned above, it is not only the feedback's design but also the personal characteristics of the feedback recipient that are important in order to explain the different effects of feedback on the learner and their learning outcomes. Several studies show

that feedback is often received sub-optimally (e.g., Anderson, Kulhavy, & Andre, 1971; Hancock, Thurman, & Hubbard, 1995; Jacobs, 2000; Jacoby, Troutman, Mazursky, & Kuss, 1984, Stark, 2001). These findings imply that conscious feedback reception is highly dependent on variables of the feedback recipient, like prior knowledge, cognitive resources, metacognitive control strategies, interest, motivation, feedback acceptance, negative emotions and personal preferences for a learning subject etc. Considering the large number of possible personal variables that can have an influence on the learner's feedback reception, we will discuss only some of these variables.

Learner's prior knowledge

The findings of Jacoby et al. (1984) show that students with a higher level of prior knowledge can implement the feedback more easily than students with lower levels of prior knowledge. Further, students with lower levels of prior knowledge need more complex forms of feedback (e.g., elaborated feedback), while for students with higher levels of prior knowledge simple forms of feedback (e.g., knowledge of result) are sufficient (Hanna, 1976).

Learner's motivation and feedback acceptance

Several studies show that in order for feedback to work, feedback information must be processed and used by the feedback recipient (Bangert-Drowns et al., 1991; Hancock, Thurman, & Hubbard, 1995). Bangert-Drowns and colleagues (1991), for example, refer to a mindful reception of feedback, whereby the individual consciously receives the feedback. This requires an adequate level of motivation (Kulhavy, 1977). Furthermore, the learner's feedback acceptance is important, meaning the credibility of feedback or to what extent the feedback is in accordance with the learner's own performance judgements or feedback expectations. For example, some studies have demonstrated that people tend to devalue feedback, when it is worse than they expected (Alloy & Abramson, 1979; Sedikides et al., 1998).

Considering the large number of possible influences on the learner's feedback reception, we will try to assess this comprehensively with the concept of educational and learning capital.

The Possible Role of Educational and Learning Capital

Before introducing the concept of educational and learning capital, we first have to describe the term 'actiotope'. Ziegler, Vialle and Wimmer (2013) define one's actiotope as comprising the individual and the material, social, and informational environment with which the individual interacts. Depending on the quality of this interplay between individual and environment, an individual delivers different performances. In order to attain any goal, resources are helpful and often necessary. In the case of learning goals or striving towards better performance levels, learning resources are required. Ziegler and Baker (2013) differentiate between exogenous and endogenous learning resources, depending on where in an actiotope the resources are situated: either in the environmental component of an actiotope (i.e., exogenous resources, also termed educational capital) or in the individual component of an actiotope (i.e., en-

dogenous resources, also termed learning capital). Furthermore, Ziegler and Baker (2013) distinguish five different forms of educational and five different forms of learning capital. For definitions, illustrations and examples of the ten capitals see Table 1.

Table 1 *Definitions, Illustrations and Sample Items of the Five Forms of Educational Capital and the Five Forms of Learning Capital*

Type of capital	Definition	Illustration	Sample item
		Educational capital	
Economic educational capital	Economic educational capital is every kind of wealth, possession, money or valuables that can be invested in the initiation and maintenance of educational and learning processes. (p. 27)	Economic educational capital can be used as a targeted support for students, to pay for stimulating educational games, for appropriate tutors and mentors, or even for the higher public transportation fee to a good university in a distant area.	My family is willing to spend more money than others for learning.
Cultural educational capital	Cultural educational capital includes value systems, thinking patterns, models and the like, which can facilitate - or hinder - the attainment of learning and educational goals. (p. 27)	Are male students who want to become a primary school teacher seen in the same light as the female students who aspire the same position? What about professor for physics: what is the common gender issue about this topic in your country?	In my social environment learning is considered to be very important.
Social educational capital	Social educational capital includes all persons and social institutions that can directly or indirectly contribute to the success of learning and educational processes. (p. 28)	Social resources can be needed in order to gain access to specific learning environments (e.g., scholarships, networks) or in order to improve/establish learning conditions (e.g., supportive partner, supportive friends).	My friends and my family support me in my learning.
Infrastructural educational capital	Infrastructural educational capital relates to materially implemented possibilities for action that permit learning and education to take place. (p. 28)	Infrastructural capital implies the availability of libraries and learning-medias in schools and universities, which can support the development of interests and facilitate good learning conditions.	I have optimal learning opportunities.
Didactic educational capital	Didactic educational capital means the assembled expertise involved in the design and improvement of educational and learning processes. (p. 29)	Didactic educational capital has increased over the last decades, due for example to better organised teaching techniques and pedagogically improved learning feedback. This is shown in rises in performance levels in all domains. For example if the IQ-Test wouldn't be continually adjusted, then the average IQ would be much higher in the recent decades (e.g., Flynn, 2007).	I use suggestions and tips on how I learn best.

	Learning capital		
Organismic learning capital	Organismic learning capital consists of the physiological and constitutional resources of a person. (p. 29)	The availability of organismic learning capital is important not only for physiological activities, but also for cognitive activities needed for study achievements (Bellisle, 2004; Gottfredson, 2004).	My very good physical condition is a good basis for my continuous learning.
Actional learning capital	Actional learning capital means the action repertoire of a person - the totality of actions they are capable of performing. (p. 30)	The current available action repertoire is a good predictor for future performance (Ziegler, 2008).	I always know exactly what I can learn.
Telic learning capital	Telic learning capital comprises the totality of a person's anticipated goal states that offer possibilities for satisfying their needs. (p. 30)	Telic learning capital is the availability of functional goals for the learning process, like planning the next learning step under good physical conditions (the importance of rest periods).	I have set myself the learning goal to learn more and more.
Episodic learning capital	Episodic learning capital concerns the simultaneous goal- and situation-relevant action patterns that are accessible to a person. (p. 31)	If a student speaks fluently English as a foreign language, this does not imply, that this student will always say the "right" thing when asked questions by the professor in an English seminar.	Due to various experiences, I know how I can achieve outstanding success.
Attentional learning capital	Attentional learning capital denotes the quantitative and qualitative attentional resources that a person can apply to learning. (p. 31)	If students spend a lot of their time on leisure activities, such as chatting with friends then the quantitative attentional resources for learning can be diminished. The attentional quality can be reduced if students do not have a quiet work place for studying at home.	I can concentrate without distractions on my learning for university.

Note. All definitions are quotes from Ziegler & Baker (2013); all sample items are from the QELC (Vladut et al., 2013);

Educational and learning capital have already been shown to be associated with high performance levels and effective learning processes (Debatin, Hopp, Vialle, & Ziegler, 2015; Harder, O'Reilly, & Debatin, 2018; Harder, Trottler, Vialle, & Ziegler, 2015; Vladut, Liu, Leana-Taɑcilar, Vialle, & Ziegler, 2013; Vladut, Vialle, & Ziegler, 2015; Ziegler & Baker, 2013; Ziegler, Debatin, & Stoeger, 2019; Ziegler & Phillipson, 2012). According to their positive role in an actiotope and in learning processes, resources should also help in dealing with feedback. Of particular interest is the case of negative feedback. For example, having a supportive environment and favorable dispositions should buffer the potentially crushing effect of negative feedback. A learner with a lot of educational and learning capital at their disposal should be able to maintain their motivation, adapt to the experience of failure affectively and through reorganizing their actions for upcoming performance situations; and, the learner should remain self-confident instead of becoming anxious because they trust their endogenous resources as well as the environment's support. By contrast, the lack of those resources should make a learner susceptible to giving up after failure, developing anxieties, and

so on. In the case of positive feedback, high levels of resources should help the learner to embrace their success and continue to perform successfully. However, this task should not be nearly as demanding as being confronted with a personal failure.

Given the width of the concept, we posit that educational and learning capital should be a valid conceptualization of the moderating circumstances, which either lead to positive or negative outcomes after feedback.

Current Study and Research Questions

The current study investigated feedback effects on (a) cognitive, (b) metacognitive, (c) motivational and (d) emotional outcomes. As not all potential variables could be examined, we focused on (a) cognitive performance after feedback, (b) coping with failure in two ways, namely adapting affective-motivationally to the failure and adapting one's actions after failure. Subsequently, we examined (c) self-confidence and stability/modifiability beliefs after feedback and (d) anxiety. We presented students with a piece of random feedback after a cognitive test they had taken one week before and also took into consideration their judgements on their pretest performance to account for the feedback's credibility. Finally, we wanted to determine the moderating role of educational and learning capital on the feedback effects as this concept captures not only characteristics of the learner that are known to moderate feedback effects, but also the resources within their environment comprehensively.

Our research questions were threefold. As the effects of feedback depend on multiple moderating variables, feedback does not necessarily lead to an improvement of performance or change motivation and emotion in a desired direction. Therefore, the first aim of our study was to analyze whether different feedback leads to different reactions in cognitive, metacognitive, motivational and emotional outcomes. In general, we assumed that the more negative the feedback experience (negative feedback combined with matching performance judgements), the more negative would be the effects on the outcome variables.

Second, given that resources are indispensable to learning processes we expected to find positive effects of high educational and learning capital on the outcome variables in general, while low capital should yield negative effects.

Third, assuming that resources play a major role in dealing with negative feedback compared to the much easier situation of accepting positive feedback, we wanted to investigate whether educational and learning capital moderate the effects of feedback on the cognitive, metacognitive, motivational, and emotional outcomes. We expected to find interaction effects stating that high educational and learning capital buffer the effects of negative feedback while in the case of positive feedback it should not appear that important.

Method

Study Design and Procedure

Our study design was an experimental field study with two measurement points and randomized assignment of participants to one of three feedback conditions. At the first time of assessment, the participants

(N = 140) were requested to provide some personal data and they completed a questionnaire on educational and learning capital. Finally, they took an abbreviated test of cognitive performance and gave their judgement on the test performance. One week later, at the second time of assessment, participants randomly received one type of feedback on their performance in the cognitive performance test, assigning them either to the *upper third* (N = 46), *middle third* (N = 51), or *lower third* (N = 43) of the participating group. Then the students were requested to complete questionnaire scales on self-confidence, coping with failure, anxiety, and implicit personality theories. Finally, the students took another cognitive performance test.

In our study we wanted to reach as many students as possible during the online semester due to the COVID 19 pandemic. Further, we wanted to eliminate person-based bias when giving feedback (e.g., unconscious non-verbal signals of the feedback giver). Thus, we chose computer-based data collection in the environment of an online lecture and a computer-based written presentation mode of the feedback.

Regarding feedback form, we chose the feedback form 'knowledge of results', because we wanted it to contain the minimum of information, in order to see how this minimal information affects the chosen variables. Regarding feedback accuracy, our given feedback was randomized, meaning the feedback could have been genuine (positive, mediocre or negative) or it could have been fake (positive, mediocre or negative). To control accuracy, we assessed students' judgements of their performance in the pretest (coded exactly as the feedback). Regarding the feedback reference norm, we chose the social reference norm, meaning the students received feedback on how good their individual performance was compared to other students. This could be seen more like an achievement situation, where anxious feelings following the feedback could arise.

Sample

Our sample consisted of N = 140 university students in teacher training, 26 male and 112 female, ranging from 19 to 38 years of age (M = 22.09, SD = 3.29) and 1 to 10 semesters of study. Most of them were in the first half of their studies (50.7 % in 2nd semester, 38.6 % in 4th semester). The mean grade of the students was calculated based on the last three grades that they had received in their studies (M = 2.38, SD = 0.57). This placed them in the upper-middle segment of the performance spectrum (the German grading system ranges from 1 *excellent* to 6 *insufficient*). ~~The~~ Participants were recruited over a mandatory psychology lecture and gave their informed consent before participating in the study.

Measures

The measures taken comprised several questionnaire scales and two cognitive performance measures, namely intelligence tests. Questionnaire scales phrased for the school setting were adapted to the university context.

Educational and Learning Capital

Educational and learning capital was measured with the QELC (Questionnaire of Educational and Learning Capital; Vladut et al., 2013), which consists of ten subscales measuring the ten forms of educational and learning capital with five items each. An ex-

ample item would be "I have optimal learning opportunities" (infrastructural learning capital). Answers were given on a six-point Likert-scale ranging from 1 *I disagree completely* to 6 *I agree completely*. We used the comprehensive scales of educational and learning capital, each consisting of the mean value over the five subscales, which yielded acceptable reliabilities of Cronbach's α = .68 for educational capital and of α = .86 for learning capital.

Self-Confidence

We used the scale "Confidence in one's own competence" developed by Dweck and Henderson (1988), later adapted by Ziegler and Stoeger (2010). The scale consists of four bipolar items presenting two opposite statements. One corresponds to a negative self-evaluation (e.g., "I am not sure that I am good enough to be successful in in my university-studies") and the other pole corresponds to a positive self-evaluation (e.g., "I am sure that I am good enough to be successful in my university-studies"). Participants answered on the aforementioned six-point Likert-scale. Reliability of the self-confidence scale was satisfactory with α = .81.

Coping with Failure

Dresel et al. (2013) developed two scales that measure the degree to which a person copes adaptively with failure or errors. Two types of reactions on errors are distinguished: the *affective-motivational adaptivity* of error reactions (e.g., "If I get the answer wrong, it spoils my good mood for the entire seminar session") and their *action adaptivity* (e.g., "When I've made a mistake, I aim to improve myself"). Each subscale comprises seven items and uses the aforementioned six-point Likert scale answering format. Reliability of the subscale affective-motivational adaptivity was α = .86, and for the subscale action adaptivity the reliability was α = .83.

Anxiety

In order to assess anxiety, a six-item scale published by Ziegler, Dresel, Schober, and Stoeger (2005) was applied. A sample item reads "when I think of my university studies, I am afraid to get a bad mark". Answers were again given on the six-point Likert-scale with a higher scale value indicating a higher anxiety level
. The reliability of the anxiety scale was good with α = .79.

Stability/Modifiability Beliefs

Dweck (e.g., 1999, 2006) proposed two different implicit personality theories, the fixed mindset assuming abilities are stable and the growth mindset assuming abilities can change. Ziegler and Stoeger (2010) refined the concept differentiating it further into stability beliefs regarding existing abilities and modifiability beliefs regarding ability deficits. They were measured with two subscales of six items each (Ziegler & Stoeger, 2010). A sample item for stability beliefs reads "After I have learned something, I don't forget how to apply it". A sample item for modifiability beliefs is "I can improve my skills". Answers were again given on the six-point Likert-scale. The reliability of the subscale stability beliefs was α = .87, and for the subscale modifiability beliefs the reliability was α = .80.

Performance

On the first measurement point, we assessed performance with a selection of 16 matrices from the RAVEM APM (Kratzmeier, 1980), which measure fluid intelligence. Due to shortening of the test (we used a selection of 16 out of the 36 matrices), no IQ scores were calculated; instead, we used raw-scores. The reliability was low with $\alpha = .60$.

On the second measurement point, we assessed performance with the BEFKI GK-C (Schipolowski, Wilhelm, & Schroeders, 2013), a test of crystalline intelligence or general knowledge. It consists of 12 multiple-choice items (one answer out of four options is correct) drawn from different domains, e.g., medicine, religion and history. Reliability of the test was lower than expected at $\alpha = .52$. We assume this was because each of the 12 questions was from a different knowledge domain.

Nevertheless, this is not highly problematic, neither for the measurement of the fluid intelligence nor for the measurement of the crystalline intelligence, because the measure of intelligence was not the objective of our current study. Hence, most scales that we used for the study showed acceptable reliabilities.

Data Analysis

To test the effects of feedback and judgements separately and combined (interaction effect) as well as the influence of educational and learning capital and its moderating effects on feedback and judgement, hierarchical regression analyses were performed with interaction terms consisting of the product term of the predictor and the moderator (see e.g., Cohen, Cohen, West, & Aiken, 2003). All models were performed by entering predictors step by step, always testing the new model against the previous version with fewer predictors (change in R^2). In a first step, the main effects of feedback and judgement were tested. Second, their interaction term was added (model 2). Both models served to evaluate our first research question. In model 3 educational and learning capital was entered into the equation to show its main effect (second research question). Finally, the interaction terms of educational and learning capital with feedback and judgement respectively were added in model 4, testing the moderating effects stated in research question number three. To obtain appropriate beta weights for the interaction terms, all variables were z-standardized beforehand and the interaction terms (cross products) were calculated with z-standardized values instead of using the z-standardized raw-variable product term (Cohen et al., 2003; LeBreton, Tonidandel, & Krasikova, 2013). All regression analyses were performed with SPSS 25 (IBM, 2017).

Moderator analysis demands hierarchically well-formulated models (Cohen et al., 2003; Hayes, 2018; Jaccard & Turrisi, 2003), i.e., all factors of a lower hierarchy order (the main effects) must be included when a higher order predictor (the interaction term) is to be analyzed. This demand can lead to collinearity among predictors which also happened in our analyses. Therefore, beta weights in regression analyses including interaction terms cannot be interpreted reliably and we will evaluate the significance of an additional predictor only by the produced change in R^2.

Unfortunately, beta weights represent the effect sizes and thereby are of great importance to interpret results (LeBreton, Hargis, Griepentrog, Oswald, & Ployhart, 2007). They allow for within model comparisons of

predictors' influence on the criterion, while the approach of interpreting changes in R^2 only allows for between model comparisons (LeBreton, Tonidandel, & Krasikova, 2013). Hence, to gain more insight into each predictor's true contribution to the total of explained variance, we also performed relative weight analyses (RWA; Johnson & LeBreton, 2004; Tonidandel & LeBreton, 2011). The relative weight (RW) refers to a predictor's true importance in explaining variance, i.e., taking into account shared variance between collinear predictors instead of neglecting those variance portions and thereby biasing the beta weights. To do so, RWA creates new orthogonal variables that are maximally related to the original predictors. Then, the dependent variable is regressed on the orthogonal variables to save their standardized regression *coefficients*. Next, the original predictors are regressed on the orthogonal variables to obtain the standardized regression *weights*. Last, the procedure multiplies each squared *weight* with each squared *coefficient*. The sum of all products an original predictor is involved in represents its RW and all RWs sum up to the total of explained variance (R^2). RWAs were carried out with the RWA-Web Tool (Tonidandel & LeBreton, 2015) which provides R Code (R Core Team, 2020). RWAs were run for each criterion on the regression model 4 including all predictors of interest to compare their relative contribution to the total explained variance. RWs were tested for significance using bias corrected and accelerated 95 % confidence intervals (equals an alpha error of 5 %) and 10,000 replications bootstrapping. In order to analyze interaction effects with RWA, LeBreton, Tonidandel and Krasikova (2013) recommend using residualized interaction terms, i.e., regressing the criterion on the predictors involved in the interaction effect (the main or lower order effects) and saving the residuals which represent the "clean" interaction effect. These residualized interaction terms represent only the higher order effect independent of the lower order effects (dependency represents an interpretation and collinearity problem in regular analyses techniques), allowing for easy interpretation and direct comparison with other effects.

Results

Descriptive Statistics and Correlations

Table 2 provides descriptive statistics for all scales as well as their bivariate correlations. Feedback and performance judgements were coded 0, 1, 2 for lower, middle and upper third of the sample respectively. Feedback was distributed about evenly across the sample as reported above (M = 1.02, SD = 0.80). Students judged their performance slightly lower and with less variance than the random feedback suggested (M = 0.95, SD = 0.60). The performance measures showed expected means: The mean pretest score was 11.22 points out of 16 (70 % correct) and for the posttest scores 8.20 points out of 12 (68 % correct). All other scales could range from 1 to 6 and most of them showed medium scores and normal standard deviations (anxiety scores were below, self-confidence, action adaptivity after failure and modifiability beliefs were above the scale center).

Correlations between all variables showed two unexpected findings. Feedback did not correlate with other variables while judgements did. Hence, the effects of random feedback alone seem to be weak, which supports the idea of additionally consider-

ing judgements to account for (in-)congruence or credibility effects. Second, the pre- and posttest performances did not correlate with other variables, which is surprising and will be discussed later. Aside from that, all variables correlated significantly in the expected direction with the exception of the non-significant relation between stability beliefs and anxiety ($r = -.15$).

Table 2 *Means, Standard Deviations, and Bivariate Correlations of All Scales*

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10
1. ELC	3.62	0.56										
2. Pretest	11.22	2.11	.041									
3. Judgement	0.95	0.60	.334**	.450**								
4. Feedback	1.02	0.80	.081	.027	−.028							
5. Self-confidence	4.20	0.88	.429**	.068	.246**	.093						
6. Affective-motivational adaptivity a. f.	3.78	0.85	.174*	−.088	−.028	.024	.301**					
7. Action adaptivity a. f.	4.33	0.64	.422**	−.000	.189*	.105	.336**	.422**				
8. Anxiety	2.63	0.90	−.228**	.031	−.028	−.097	−.568**	−.507**	−.207*			
9. Stability beliefs	3.33	0.82	.227**	.065	.104	.061	.257**	.299**	.256**	−.150		
10. Modifiability beliefs	4.74	0.63	.316**	−.122	.154	.081	.546**	.483**	.522**	−.408**	.230**	
11. Posttest	8.20	2.09	.049	.138	.065	.023	.047	.025	.016	−.063	.115	−.009

Note. * $p < .05$, ** $p < .01$, ELC = educational and learning capital, a. f. = after failure;

Effects of Feedback and Performance Judgement

Models 1 and 2 tested the influence of feedback and performance judgements. As can be seen in table 3, feedback did not show any significant effects on response variables while students' judgements influenced their self-confidence and action adaptivity after failure, meaning higher judgements came with higher self-confidence and higher action adaptivity after failure.

The interaction term feedback x judgement raised the R^2 in models 2 slightly for action adaptivity ($\Delta R^2 = .025$, $p < .10$) and significantly for affective-motivational adaptivity after failure ($\Delta R^2 = .057$, $p < .01$). Figures 1 and 2 visualize the effects. For affective-motivational adaptivity we find a congruence effect, meaning that the highest adaptivity scores coincide with matching judgements and feedback. For mediocre judgements the differences appear to be very small, so all sorts of feedback seem to be easily acceptable, whereas high and low judgements combined with incongruent feedback lead to lower affective-motivational adaptivity. For action adaptivity (figure 2) we find the same pattern but less pronounced (as expected by the marginally significant effect). Hence, congruence played a role for adapting after failure but not for other outcomes.

Direct Effects of Educational and Learning Capital

Models 3 (see table 3) added educational and learning capital as a predictor to the regressions. We find significant gains in explained variance compared to model 2 without this predictor for all variables, except the posttest which remained unaffected by educational and learning capital. The signif-

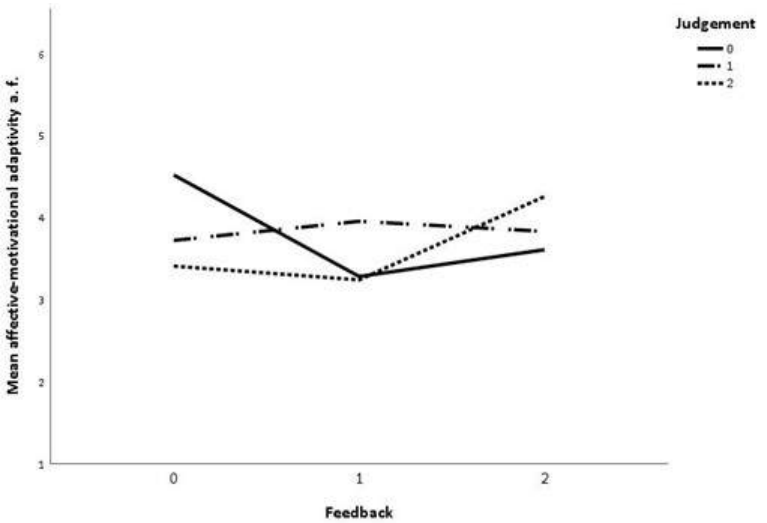


Figure 1 Interaction of feedback and performance judgement on students' affective-motivational adaptivity after failure.

icant changes in explained variance range from ΔR^2 = .027 for affective-motivational adaptivity after failure up to ΔR^2 = .123 and ΔR^2 = .125 for self-confidence and action adaptivity after failure. Relying on the signs of the beta weights these effects are all positive, i.e., higher capital coming with more favorable motivational and affective responses (also less anxiety, the negative sign here refers to the desired association).

Moderating Effects of Educational and Learning Capital

A moderating effect of educational and learning capital was found for anxiety and stability beliefs (ΔR^2 = .052, p < .05 and ΔR^2 = .068, p < .01). In the case of anxiety this is based on the interaction of educational and learning capital and judgement with a RW = .052 (equals the ΔR^2), in the case of stability beliefs on the other hand, the interaction of educational and learning capital with feedback contributes RW = .050 to the R^2.

In both cases, figures 3 and 4 indicate that the effect stems from students with low educational and learning capital. For anxiety (figure 3), only these students show different anxiety levels depending on their performance judgements. Thinking they have done well in the pretest led to more anxiety than judging their pretest performance as mediocre and they seemed to experience even less anxiety when having judged their performance as low. Hence, the higher their judgements, the higher their anxiety when they possessed low educational and learning capital.

For stability beliefs among students with low educational and learning capital, figure 4 indicates that more negative feedback led to lower stability beliefs, i.e., these students lost some faith in the stability of their competence after one piece of nega-

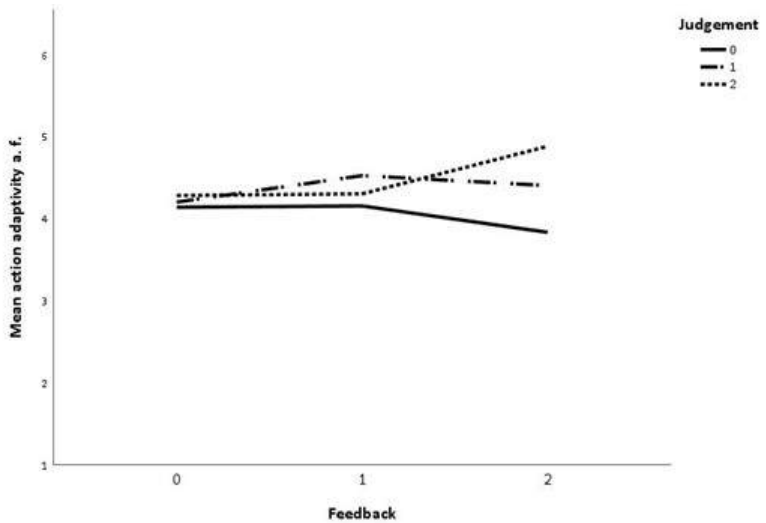


Figure 2 Interaction of feedback and performance judgement on students' action adaptivity after failure.

Table 3 *Results of Hierarchical Regression Analyses and Relative Weight Analyses for All Outcome Variables*

Predictor	Self–confidence					Action adaptivity a. f.					Affective–motivational adaptivity a. f.					Anxiety				
	β	ΔR^2	R^2	RW[b]	RS–RW (%)[c]	β	ΔR^2	R^2	RW[b]	RS–RW (%)	β	ΔR^2	R^2	RW[b]	RS–RW (%)	β	ΔR^2	R^2	RW[b]	RS–RW (%)
Model 1:		.070	.070**				.048*	.048*				.001	.001				.010	.010		
Feedback	.100					.111					.024					−.098				
Judgement	.249***					.192*					−.028					−.031				
Model 2:		.008	.079*				.025+	.073*				.057**	.059*				.008	.019		
Constant	.003					.005					.007					−.003				
Feedback	.092					.097					.003					−.090				
Judgement	.261***					.213*					.004					−.043				
Feedback x Judgement	.095					.166+					.251**					−.095				
Model 3:		.123***	.202***				.125***	.198***				.027*	.085*				.046*	.064+		
Constant	.001					.003					.006					−.002				
Feedback	.062					.067					−.011					−.072				
Judgement	.128					.078					−.058					.038				
Feedback x Judgement	.046					.116					.228**					−.065				
ELC	.377***					.381***					.175*					−.230*				
Model 4:		.012	.214***				.015	.214***				.025	.110*				.052*	.116*		
Constant	−.037					−.031					−.048					.077				
Feedback	.052			.006	2.6	.057			.007	3.2	−.025			.000	0.2	−.051			.006	4.9
Judgement	.116			.037	17.3	.078			.022	10.4	−.071			.003	2.9	.057			.002	1.4
Feedback x Judgement[a]	.018			.002	1.2	.133			.017	8.0	.204*			.045	40.5	−.035			.004	3.6
ELC	.388***			.155*	72.5	.400***			.152*	71.1	.195*			.030	27.0	−.259**			.050	43.0
ELC x Feedback[a]	.051			.003	1.4	−.055			.001	0.6	.035			.008	6.9	−.039			.003	2.1
ELC x Judgement[a]	.100			.011	5.0	.117			.014	6.7	.154+			.025	22.5	−.227**			.052	45.0

Note. Interaction terms represent the cross–product of the z–standardized variables instead of the standardization of the raw–variable cross–product (Cohen et al., 2003). In consequence, their mean will most likely deviate from zero which is why the constant/intercept is reintroduced in models including interaction terms (LeBreton et al., 2013). [a] Relative weight analysis used the residuals of the interaction terms regressed on their constituting variables, yielding a "clean" interaction effect (LeBreton et al., 2013). [b] RW = Relative Weight (they sum up to R^2), the RWA only tested for $p < .05$; [c] RS–RW (%) = Rescaled Relative Weight (RS divided by R^2, contribution of each predictor to the R^2 in percent). + $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$, ELC = educational and learning capital, a. f = after failure.;

Table 3 (continued)

Predictor	Stability beliefs					Modifiability beliefs					Posttest				
	β	ΔR^2	R^2	RW[b]	RS−RW (%)[c]	β	ΔR^2	R^2	RW[b]	RS−RW (%)	β	ΔR^2	R^2	RW[b]	RS−RW(%)
Model 1:		.015	.015				.031	.031				.005	.005		
Feedback	.063					.085					.025				
Judgement	.106					.156+					.066				
Model 2:		.005	.020				.011	.042				.005	.010		
Constant	.002					.003					−.002				
Feedback	.057					.076					.031				
Judgement	.115					.170*					.056				
Feedback x Judgement	.075					.111					−.075				
Model 3:		.036*	.056+				.069*	.111**				.001	.011		
Constant	.001					.002					−.002				
Feedback	.041					.054					.028				
Judgement	.043					.071					.042				
Feedback x Judgement	.048					.074					−.080				
ELC	.205*					.282**					.039				
Model 4:		.068**	.125**				.028	.139**				.000	.012		
Constant	−.007					−.049					−.010				
Feedback	.036			.002	1.6	.040			.004	2.7	.026			.001	5.2
Judgement	.071			.008	6.8	.066			.015	10.7	.040			.003	23.8
Feedback x Judgement[a]	.162+			.013	10.1	.084			.007	5.2	−.083			.006	47.7
ELC	.235**			.042	33.4	.307**			.084*	60.8	.042			.002	16.2
ELC x Feedback[a]	−.279**			.050	40.4	−.046			.001	0.6	.003			.000	3.3
ELC x Judgement[a]	.102			.010	7.7	.166*			.028	20.0	.022			.000	3.8

tive feedback. For mediocre feedback the effect seems to be less pronounced and for positive feedback the students with low educational and learning capital seem to react similarly to their peers with higher educational and learning capital.

Seeing the Whole Picture: Comparing Effect Sizes of Predictors

Summarizing these results, we found scattered effects of feedback and judgements: main effects in two out of seven outcome variables (i.e., self-confidence and action adaptivity after failure), and interaction or congruence effects in two outcomes (action and affective-motivational adaptivity after failure). Main effects of educational and learning capital emerged in all outcome variables except the posttest and interaction or moderating effects of educational and learning capital on the effect of feedback/judgement emerged again in two of the outcomes (anxiety and stability beliefs). These seemingly inconsistent findings based on significant gains in R^2 should be viewed in the context of all predictors and their true effect sizes to see the whole picture. The RWAs performed on the complete regression models (model 4) provide us with each predictors' true importance for the total explained variance (see table 3).

First and foremost, for all but one outcome the total R^2 reached significance, ranging between 11.0 % (affective-motivational adaptivity after failure) and 21.4 % (self-confidence and action adaptivity after failure). The posttest score could not be predicted solidly (R^2 = 1.2 %).

Taking a look at the rescaled relative weights (RS-RWs) of the predictable outcomes we can identify two patterns: outcomes that heavily depend on educational and learning capital and outcomes that de-

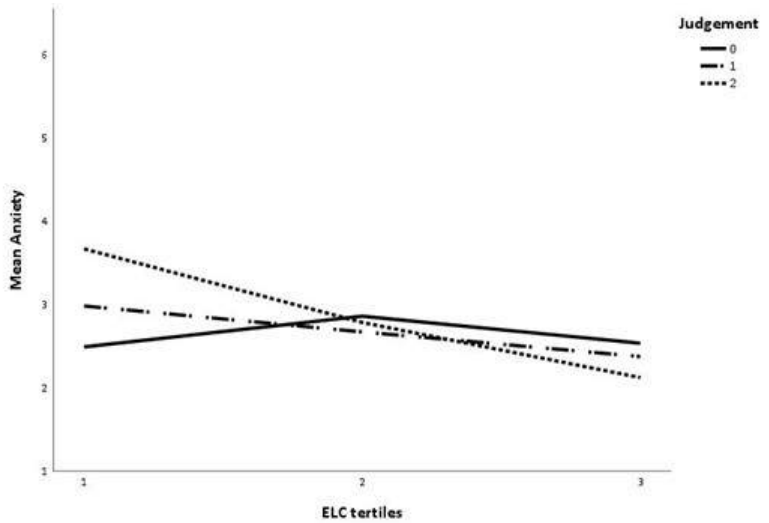


Figure 3 Interaction effect of educational and learning capital (ELC) and students' performance judgement on anxiety (to ensure the figure's readability ELC was classified into tertiles only for this visualization).

pend on several predictors, very often the interaction terms.

Self-confidence, action adaptivity after failure and modifiability beliefs depended strongly on educational and learning capital, which explained 60.8 % to 72.5 % of the respective total R^2. For Self-confidence the second best predictor was judgment (RS-RW = 17.3 %); for action adaptivity after failure, judgement and the congruence of judgement and feedback contributed 10.4 % and 8.0 % to the total R^2; and for modifiability beliefs the interaction of educational and learning capital with judgment (RS-RW = 20.0 %) and judgement alone (RS-RW = 10.7 %) contributed considerably to the total R^2.

The other three predictable outcomes did not show a clear main predictor. Affective-motivational adaptivity after failure was largely explained by the congruence of feedback and judgement (RS-RW = 40.5 %), by educational and learning capital (RS-RW = 27.0 %) and also by the interaction of educational and learning capital and judgment (RS-RW = 22.5 %).

Anxiety depended on two predictors equally, namely the interaction of educational and learning capital and judgment with RS-RW = 45.0 % and the educational and learning capital itself with RS-RW = 43.0 %.

Last, stability beliefs depended on almost all available predictors (only feedback fell behind the other predictors' contribution). Two predictors were more pronounced (interaction of educational and learning capital and feedback, RS-RW = 40.4 %, and educational and learning capital itself, RS-RW = 33.4 %), while the congruence of feedback and judgment (RS-RW = 10.1 %), the interaction of educational and learning capital and judgement (RS-RW = 7.7 %), and judgment (RS-RW = 6.8 %) contributed similar smaller amounts of explained variance.

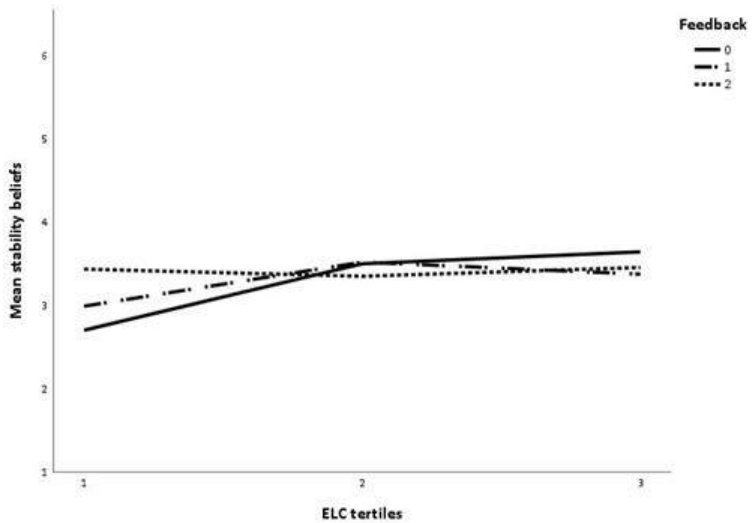


Figure 4 Interaction effect of educational and learning capital (ELC) and feedback on students' stability beliefs (to ensure the figure's readability ELC was classified into tertiles only for this visualization).

Discussion

This study investigated (a) the effects of random performance feedback and the learners' judgement of their performance on cognitive, metacognitive, motivational and emotional outcomes as well as (b) the supposedly positive effect of having educational and learning capital at one's disposal. In addition, we analyzed whether (c) educational and learning capital moderates the feedback and judgement effects, assuming that disposing over educational and learning capital would buffer negative consequences after experiences of failure (receiving negative feedback). We found that feedback and judgements had rather weak effects whereas educational and learning capital directly affected most responses. A moderating effect of educational and learning capital was found for some responses but not for all.

Overall, it is encouraging that almost all outcomes could be well predicted by feedback, judgements, educational and learning capital and their interaction effects bespeaking the selection of predictors, especially the validity of educational and learning capital. However, the combinations of relevant predictors also introduced new questions. We will discuss each research question in detail in the following.

Effects of Feedback and Performance Judgements

Our first research question treated the effects of feedback in general and, to account for the credibility of our random feedback, the effect of students' performance judgements and how they matched the received random feedback (interaction effect).

Surprisingly, we found no effects of feedback alone on any outcome. This can mean either that students were aware of the feedback but not affected by it or – more likely – that students did not effectively receive the feedback. Due to the available possibilities (online lectures during the COVID19 pandemic), our study setting was not perfect in terms of guaranteeing feedback reception. In order for feedback to work, the learner must process the feedback information. We used computer-based feedback which we assume requires a higher motivation of the recipient to think about it intensively than a person-based feedback does (Krause, 2007). It was also delayed by one week while some studies indicate that immediate feedback works better (Bangert-Drowns et al., 1991; Kulik & Kulik, 1988). Furthermore, we only gave feedback based on the social reference norm (e.g., "Your answers place your performance in the upper third of the performance spectrum of this student group"). More performance information like adding the percentage or total number of correct solutions (criterial norm) or giving elaborative information might have lent more relevance to the feedback (Collins, Carnine, & Gersten, 1987; Kluger & DeNisi, 1996; Kopp & Mandl, 2014).

Performance judgements on the other hand displayed two effects, namely on self-confidence and action adaptivity after failure. Judging their own performance positively made students more self-confident, which concurs with theories and findings on the development of the academic self-concept through personal experience (Marsh & Shavelson, 1985; Shavelson, Hubner, & Stanton, 1976). They also estimated their action adaptivity after failure more positively after having a positive performance experience. Possibly, experiencing success made them optimistic dealing with future learning situations.

We did not explicitly expect direct judgement effects as we were mostly interested in the interaction effect of feedback and performance judgements representing the congruence between both, i.e., the credibility of the feedback. This credibility aspect showed effects on both coping with failure measures: one clear effect for affective-motivational adaptivity after failure (explaining 40.5 % of the total R^2 in the RWA) and a marginally significant effect for action adaptivity after failure. Congruent negative and positive feedback led to better affective-motivational adaptivity. For action adaptivity the effect seemed present only for congruent positive feedback. Obviously, credible feedback, even when negative, can enhance coping with failure. The question remains, why none of the other outcomes depended on the credibility of feedback. For example, Shibata et al. (2009) demonstrated that in the case of perceptual learning even fake feedback modulated learning (fake vs. genuine feedback condition). Here only a few students questioned the credibility of feedback (17 out of 84), and only five out of the 17 students were aware of the discrepancy between the received and their expected feedback (free report), but two of these five students were in the genuine feedback condition. Hence, participants seem to be rather unreliable when it comes to judge the accuracy of feedback which might explain the lack of effects on other responses in our study. However, a look at existing evidence shows that the credibility of feedback is rarely considered and more research could shed light on these effects.

Effects of Educational and Learning Capital on Feedback Responses

Our second research question treated the direct effects of having educational and learning capital at one's disposal. Our analyses confirmed the close relatedness of educational and learning capital with reactions in a learning process (e.g., Vladut et al., 2013; Vladut, Vialle, & Ziegler, 2015; Ziegler, Chandler, Vialle, & Stoeger, 2017; Ziegler, Debatin, & Stoeger, 2019). In all cases, the effect of educational and learning capital was positive, hence leading to desirable metacognitive, motivational, and affective responses (cognitive performance was not predictable by any predictors, see below).

This expected direct effect of educational and learning capital was especially pronounced in three outcomes. It predominantly predicted self-confidence, modifiability beliefs and action adaptivity after failure, contributing more than 60 % to the total explained variance. This is largely in line with the findings of Vladut, Vialle and Ziegler (2016) where educational and learning capital explained 9 to 34 % of the variances of self-confidence, stability/modifiability beliefs and coping with failure (combined scale of affective-motivational adaptivity and action adaptivity after failure). If we look at a failed learning process those three variables can be regarded as key elements to deal with failure: trusting their abilities helps learners to continue their efforts, modifiability beliefs also favor taking action to eliminate deficits and actually adapt their actions after failure to ensure more successful learning in the future (cf. Dweck, 1999). Educational and learning capital thus could be viewed as a booster of key elements for successful learning careers which inevitably will provide experiences of

failure at some point. At the same time, it means that educational and learning capital would not be that important for the emotional aspects of adapting after failure and anxiety in the learning context as well as for stability beliefs, which is harder to explain. As this is all speculation, future studies need to shed light on the explanation behind this finding.

Finally, the only outcome that could not be predicted by any predictor was cognitive performance in the posttest (R^2 = 1.2 %). As we already noticed in table 2 it did not correlate with any other variable which points to a problem with the test itself. It was a test on crystalline intelligence (BEFKI GC-K, Schipolowski, Wilhelm, & Schroeders, 2013) comprising questions on general knowledge in very diverse domains. According to the test manual, the BEFKI GC-K has been developed in order to economically measure crystalline intelligence. It correlates for example with self-reported knowledge (r = .52) and it is reliable (α = .81; Schipolowski, Wilhelm, & Schroeders, 2013). Unfortunately, this has not been the case in our study. Maybe students mainly aged 19 to 25 no longer deem these areas of knowledge relevant (sample item: "Families- and inheritance law are a subject matter of the a) Civil Code, b) Social Security Code, c) Basic Law, or d) Community Code"), or participants were no longer motivated to engage in this last assessment after having filled out the questionnaires beforehand.

Moderating Effects of Educational and Learning Capital

Our third research question was concerned with educational and learning capital as a moderator to the effects of feedback, especially in the case of (credible) negative feedback. We expected high educational and learning capital to buffer negative consequences of a failure experience. Conversely, low educational and learning capital should result in more detrimental consequences for the learner.

Indeed, we found two moderating effects pointing in this direction as they both originated from students with low educational and learning capital and showed negative consequences in terms of higher anxiety and lower stability beliefs. The first one was an interaction of educational and learning capital and judgements, suggesting that the higher the judgements, the higher students' anxiety when they possessed low educational and learning capital. Judging their performance as very good might lead to high expectations for future performance. At the same time, those students know that they possessed low educational and learning capital which might make them anxious to fail. Similar lines of thinking and feeling are known from students who identify a learning situation as threatening for their self-concept and self-esteem, and thereby experience high levels of anxiety. In consequence, these students are less likely to learn (Weiss, 2000). Getting hopes up by a positive self-judgement might exactly trigger this process. On the other hand, as soon as educational and learning capital was at a medium to high level, the detrimental effect disappeared or was indeed buffered.

The second moderating effect found, was educational and learning capital influencing how feedback affects stability beliefs. In particular, negative feedback in combination with low educational and learning capital led to lower stability beliefs. These students' beliefs in the stability of their competence seemed to be more fragile in the face of failure due to their unfavorable

levels of resources. If we replace 'resources' with examples subsumed under the concept of educational and learning capital this finding is not surprising. For example, the presence of social support is an important factor associated with student success (e.g., Dao, Lee, & Chang, 2007; Laurence, Williams, & Eiland, 2009). Again, having more educational and learning capital made the unfavorable consequences disappear, as we expected.

Hence, the question that remains is why we did not find moderating effects on other outcome variables. For example, why is their self-confidence not shaken while stability beliefs appear fragile? Why is their coping with failure unaffected by the moderator? Given the data we have at hand, the only explanation we can provide is that negative consequences seem to occur under extreme circumstances, i.e., when negative feedback or high hopes (judgments) coincide with low educational and learning capital; in the middle to positive spectrum of student characteristics we did not observe detrimental effects. As combinations of extreme conditions constitute only a small part of the sample, our study design might be unable to reveal the interaction effects. Instead, one would need to investigate only students with low educational capital or confront participants with only negative feedback, and so on.

Limitations and Outlook

One limitation of the current study is the lack of detectable feedback effects. We discussed possible reasons and suggest that future studies carefully select their feedback design to enhance its effect on students (e.g., person-based delivery, ensuring processing, not random feedback). Additionally, it is advisable to check students' awareness of feedback (in-)accuracy (e.g., Shibata et al., 2009) in case of working with random feedback. We assessed students' judgements, contrasted them with the random feedback received and assumed from the congruence or incongruence that students believed the feedback or may have found it incredible. In fact, we do not know if students believed or disbelieved incongruent feedback or even were aware of the possible inaccuracy. Therefore, it would be useful to control whether participants notice discrepancies between their expected and received feedback.

A second implication for future studies arises from the questions emerging from this study. We partly confirmed our moderating hypothesis but cannot tell for sure why other variables did not display the expected effects. It might help to expose interaction effects under extreme conditions by investigating participants with extreme characteristics (underprivileged learning environments, learning disorders, etc.) or creating extreme conditions through the study design (cf. negative feedback experiments by Mueller and Dweck, 1998).

In a similar vein, we can only speculate why the credibility of feedback only prove relevant for coping with failure and why educational and learning capital predominantly predicted three outcomes, while three other outcomes showed quite different patterns of predictor relevancy. As our combination of predictors was not used before on these outcomes we cannot draw any inferences from these findings but need clarification from further research.

References

Alloy, L. B., & Abramson, L. Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology, 108*(4), 441–485.

Anderson, R. C., Kulhavy, R. W., & Andre, T. (1971). Feedback procedures in programmed instruction. *Journal of Educational Psychology, 62,* 148–156.

Balzer, W. K., Doherty, M. E., & O'Connor, R., Jr. (1989). Effects of cognitive feedback on performance. *Psychological Bulletin, 106*(3), 410–433. https://doi.org/10.1037/0033-2909.106.3.410

Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A., & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61,* 213–238.

Bellisle, F. (2004). Effects of diet on behavior and cognition in children. *British Journal of Nutrition, 92*, 227–232.

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65,* 245–281. https://doi.org/10.3102/00346543065003245

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Lawrence Erlbaum Associates.

Collins, M., Carnine, D., & Gersten, R. (1987). Elaborated corrective feedback and the acquisition of reasoning skills: A study of computer-assisted instruction. *Exceptional Children, 54,* 254–262.

Dao, T. K., Lee, D., & Chang, H. L. (2007). Acculturation level, perceived English fluency, perceived social support level, and depression among Taiwanese international students. *College Student Journal, 41*(2), 287–295.

Debatin, T., Hopp, M., Vialle, W., & Ziegler, A. (2015). Why experts can do what they do: the effects of exogenous resources on the Domain Impact Level of Activities (DILA). *Psychological Test and Assessment Modeling, 57*(1), 94–110. https://doi.org/10.1037/0012-. 1649.43.1.159

Deci, E. L., Koestner, R., & Ryan, R. M. (2001). Extrinsic rewards and intrinsic motivation in education: Reconsidered once again. *Review of Educational Research, 71*(1), 1–27.

Dempsey, J., Driscoll, M. P., & Swindell, L. K. (1993). Text-based feedback. In J. Dempsey & G. Sales (Eds.), *Interactive instruction and feedback* (pp. 21–54). Educational Technology.

Dresel, M., Schober, B., Ziegler, A., Grassinger, R., & Steuer, G. (2013). Affektiv-motivational adaptive und handlungsadaptive Reaktionen auf Fehler im Lernprozess [Affective-motivational adaptive and action adaptive reactions on errors in learning processes]. *Zeitschrift für Pädagogische Psychologie, 27,* 255–271. https://doi.org/10.1024/1010-0652/a000111

Dresel, M., & Ziegler, A. (2006). Langfristige Förderung von Fähigkeitsselbstkonzept und impliziter Fähigkeitstheorie durch computerbasiertes attributionales Feedback [Long-term promotion of academic self-concept and implicit skill theory through computer-based attributional feedback]. *Zeitschrift für Pädagogische Psychologie, 20*(1/2), 49–63.

Dweck, C. S., & Henderson, V. L. (1988). *Theories of intelligence: Background and measures.* University of Illinois. Unpublished manuscript.

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality and development.* Psychology Press.

Dweck, C. S. (2006). *Mindset: The new psychology of success.* Random House.

Flynn, J. R. (2007*). What is intelligence? Beyond the Flynn effect.* Cambridge University Press.

Fong, C. J., Patall, E. A., Vasquez, A. C., & Stautberg, S. (2018). A meta-analysis of negative feedback on intrinsic motivation. *Educational Psychology Review, 31,* 121–162. https://doi.org/10.1007/s10648-018-9446-6

Geurten, M., & Meulemans, T. (2016). The effect of feedback on children's metacognitive judgments: a heuristic account. *Journal of Cognitive Psychology, 29*(2), 184–201.

Gottfredson, L. S. (2004). Life, death and intelligence. *Journal of Cognitive Education and Psychology, 4,* 23–46.

Gist, E. M., & Mitchell, T. R. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. *The Academy of Management Review, 17,* 183–211. https://doi.org/10.5465/AMR.1992.4279530

Hanna, G. S. (1976). Effects of total and partial feedback in multiple-choice testing upon learning. *Journal of Educational Research, 69*(5), 202–205.

Hancock, T. E., Thurman, R. A., & Hubbard, D. C. (1995). An expanded control model for the use of instructional feedback. *Contemporary Educational Psychology, 20,* 410–425.

Harder, B., Trottler, S., Vialle, W., & Ziegler, A. (2015). Diagnosing resources for effective learning via teacher and parent checklists. *Psychological Test and Assessment Modeling, 57*(2), 201–221.

Harder, B., O'Reilly, C., & Debatin, T. (2018). Intelligence, educational and learning capital, and domain impact level of activities as predictors of school achievement. *Journal for the Education of the Gifted, 41*(4), 327–347. https://doi.org/10.1177/0162353218799440

Harks, B., Rakoczy, K., Hattie, J., Besser, M., & Klieme, E. (2014). The effects of feedback on achievement, interest and self-evaluation: The role of feedback's perceived usefulness. *Educational Psychology, 34*(3), 269–290. https://doi.org/10.1080/01443410.2013.785384

Hattie, J. A. C. (2009). *Visible Learning. A synthesis of over 800 meta-analyses relating on achievement.* Routledge.

Hayes, A. F. (2018). *Mediation, moderation, and conditional process analysis* (2nd ed.). Guilford Press.

IBM. (2017). *SPSS Statistics 25, Version 25.0.1 [Computer program].* IBM Deutschland GmbH.

Jaccard, J., & Turrisi, R. (2003). *Interaction effects in multiple regression* (2nd ed.). Sage.

Jacobs, B. (2000). *Feedback nach Bearbeitung einer Probeklausur – wird es genutzt? [Feedback after processing a mock exam—is it used?].* Bildungswissenschaften der Universität des Saarlandes.

Jacobs, B. (2002). *Aufgaben stellen und Feedback geben [Set tasks and give feedback].* Medienzentrum der Philosophischen Fakultät der Universität Saarbrücken.

Jacoby, J., Troutman, T., Mazursky, D., & Kuss, A. (1984). When feedback is ignored: Disutility of outcome feedback. *Journal of Applied Psychology, 69*(3), 531–545. https://doi.org/10.1037/0021-9010.69.3.531

Johnson, J. W., & LeBreton, J. M. (2004). History and use of relative importance indices in organizational research. *Organizational Research Methods, 7*(3), 238–257. https://doi.org/10.1177/1094428104266510

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254–284. https://doi.org/10.1037/0033-2909.119.2.254

Kopp, B., & Mandl, H. (2014). Lerntheoretische Grundlagen von Rückmeldungen [Learning theoretical basics of feedback]. In H. Ditton, & A. Müller (Eds.), *Feedback und Rückmeldungen. Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder [Feedback. Theoretical basics, empirical findings, practical fields of application]* (pp. 29–41). Waxmann.

Kratzmeier, H. (Ed.). (1980). *RAVEN-Matrizen-Test. Advanced Progressive Matrices.* Von J. C. Raven. Deutsche Bearbeitung von H. Kratzmeier unter Mitarbeit von R. Horn. Belz Test Gesellschaft.

Krause, U.-M. (2007). *Feedback und kooperatives Lernen [Feedback and cooperative learning].* Waxmann.

Krause, U.-M., & Stark, R. (2004). Too much of a good thing? Unwanted side effects of successful instructional interventions. In P. Gerjets, P. A. Kirschner, J. Elen, & R. Joiner (Eds.), *Instructional design for effective and enjoyable computersupported learning* (CD-ROM). Knowledge Media Research Center. https://www.iwm-tuebingen.de/workshops/SIM2004/pdf_files/Krause_et_al.pdf

Kulhavy, R. W. (1977). Feedback in written instruction. *Review of Educational Research, 47*(1), 211–222.

Kulhavy, R. W., & Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review, 1*(4), 279–308. https://www.jstor.org/stable/23359222

Kulik, J. A., & Kulik, C.-L. C. (1988). Timing of feedback and verbal learning. *Review of Educational Research, 58*(1), 79–97.

Laurence, B., Williams, C., & Eiland, D. (2009). Depressive symptoms, stress, and social support among dental students at a historically black college and university. *Journal Of American College Health, 58*(1), 56–63.

LeBreton, J. M., Hargis, M. B., Griepentrog, B., Oswald, F. L., & Ployhart, R. E. (2007). A multidimensional approach for evaluating variables in organizational research and practice. *Personnel Psychology, 60*, 475–498.

LeBreton, J. M., Tonidandel, S., & Krasikova, D. V. (2013). Residualized Relative Importance Analysis. *Organizational Research Methods, 16*(3), 449–473. https://doi.org/10.1177/1094428113481065

London, M. (2002). *Leadership development: Paths to self-insight and professional growth.* Lawrence Erlbaum Associates Publishers.

London, M. (2003). *Job feedback: giving, seeking, and using feedback for performance improvement* (2nd ed.). Erlbaum.

Mantzicopoulos, P. (1990). Coping with school failure: Characteristics of students employing successful and unsuccessful coping strategies. *Psychology in the Schools, 27,* 138–143. https://doi.org/10.1002/15206807(199004)27:2<138:AID-PITS2310270208>3.0.CO;2-8

Marsh, H. W., & Shavelson, R. J. (1985). Self-concept: Its multifaceted, hierarchical structure. *Educational Psychologist, 20*(3), 107–123. https://doi:10.1207/s15326985ep2003_1

Molin, F., Haelermans, C., Cabus, S., & Groot, W. (2020). The effect of feedback on metacognition. A randomizied experiment using polling technology. *Computers & Education, 152,* 103885. https://doi.org/10.1016/j.compedu.2020.103885

Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology, 75*(1), 33–52. https://doi.org/10.1037/0022-3514.75.1.33

Mietzel, G. (2005). *Wege in die Psychologie [Ways to psychology]* (12th ed.). Klett-Cotta.

Newton, T. J., & Keenan, A. (1985). Coping with work-related stress. *Human Relations, 38,* 107–126. https://doi.org/10.1177/001872678503800202

O'Leary, K., & O' Leary, S. (1977). *Classroom management: The successful use of behavior modification.* Pergamon.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66,* 543–578. https://doi.org/10.3102/00346543066004543

Pekrun, R., Elliot, A. J., & Maier, M. A. (2006). Achievement goals and discrete achievement emotions: A theoretical model and prospective test. *Journal of Educational Psychology, 98*(3), 583–597. https://doi.org/10.1037/0022-0663.98.3.583

Petersen, L.-E., Stahlberg, D., & Dauenheimer, D. (2000). Selbstkonsistenz und Selbstwerterhöhung: Der integrative Selbstschemaansatz [Self-consistency and self-esteem enhancement: The integrative self-scheme approach]. In W. Greve (Ed.), *Psychologie des Selbst [Psychology of the self]* (pp. 227–238). Beltz PsychologieVerlagsUnion. https://www.researchgate.net/profile/Lars-Eric_Petersen/publication/274303131_Selbstkonsistenz_und_Selbstwerterhohung_Der_Integrative_Selbstschemaansatz/links/5550b30308ae93634ec9ebe3/Selbstkonsistenz-und-Selbstwerterhoehung-Der-Integrative-Selbstschemaansatz.pdf

R Core Team (2020). *R: a language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org/

Sansone, C. (1986). A question of competence: The effects of competence and task feedback on intrinsic interest. *Journal of Personality and Social Psychology, 51*(5), 918–931. https://sci-hub.st/10.1037/0022-3514.51.5.918

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research, 46*(3), 407–441. https://www.jstor.org/stable/1170010

Schipolowski, S., Wilhelm, O., & Schroeders, U. (2013). BEFKI GC-K. Berliner Test zur Erfassung fluider und kristalliner Intelligenz – GC-Kurzskala [BEFKI GC-K. Berlin test for recording of fluid and crystalline intelligence - GC-short scale]. In C. J. Kemper, E. Brähler & M. Zenger (Eds.), *Psychologische und sozialwissenschaftliche Kurzskalen – Standardisierte Erhebungsinstrumente für Wissenschaft und Praxis* (pp. 30–34). Medizinisch Wissenschaftliche Verlagsgesellschaft.

Sedikides, C., Campbell, W. K., Reeder, G. D., & Elliot, A. J. (1998). *The self-serving bias in relational context. Journal of Personality and Social Psychological, 74,* 378–386. https://doi.org/10.1037/0022-3514.74.2.378

Sedikides, C., & Strube, M. J. (1995). The multiply motivated self. *Personality and Social Psychology Bulletin, 21*(12), 1330–1335. https://doi.org/10.1177%2F01461672952112010

Shibata, K., Yamagishi, N., Ishii, S., & Kawato, M. (2009). Boosting perceptual learning by fake feedback. *Vision Research, 49,* 2574–2585.

Stark, R. (2001). *Analyse und Förderung beispielbasierten Lernens: Anwendung eines integrativen Forschungsparadigmas [Analysis and promotion of example-based learning: Application of an integrative research paradigm].* Unveröff. Habilitationsschrift, Ludwig-Maximilians-Universität, Lehrstuhl für Empirische Pädagogik und Pädagogische Psychologie.

Swann, W., Jr. (2011). Self-verification theory. In P. V. Lang, A. Kruglanski, & E. T. Higgins (Eds.), *Handbook of Theories of Social Psychology* (pp. 23–42). Sage.

Tennant, M. (1997). Psychology and adult learning. Routledge.

Tonidandel, S., & LeBreton, J. M. (2011). Relative Importance Analysis: A Useful Supplement to Regression Analysis. *Journal of Business and Psychology, 26*(1), 1–9. https://doi.org/10.1007/s10869-010-9204-3

Tonidandel, S., & LeBreton, J. M. (2015). RWA Web: A free, comprehensive, web-based, and user-friendly tool for relative weight analyses. *Journal of Business and Psychology, 30*(2), 207–216. https://doi.org/10.1007/s10869-014-9351-z

Värlander, S. (2008). The role of students' emotions in formal feedback situations. *Critical Perspectives, 13*(2), 145–156.

Vladut, A., Liu, Q., Leana-Tascilar, M. Z., Vialle, W., & Ziegler, A. (2013). A cross-cultural validation study of the Questionnaire of Educational and Learning Capital (QELC) in China, Germany and Turkey. *Psychological Test and Assessment Modeling, 55*(4), 462–478.

Vladut, A., Vialle, W., & Ziegler, A. (2015). Learning resources within the Actiotope: A validation study of the QELC (Questionnaire of Educational and Learning Capital). *Psychological Test and Assessment Modeling, 57*(1), 40–56.

Vladut, A., Vialle, W., & Ziegler, A. (2016). Two studies of the empirical basis of two learning resource-orientated motivational strategies for gifted educators. *High Ability Studies, 27*(1), 39–60. https://doi.org/10.1080/13598139.2016.1212702

Vollmeyer, R., & Rheinberg, F. (2005). A surprising effect of feedback on learning. *Learning and Instruction, 15,* 589–602.

Weiner, B. (1986). *An attributional theory of motivation and emotion.* Springer.

Weiss, R. P. (2000). Emotion and learning – Implications of new neurological research of training techniques. *Training and Development, 54*(11), 44–48.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit formation. *Journal of Comparative Neurology and Psychology*, *18*, 459–482.

Ziegler, A. (2008). *Hochbegabung [Giftedness]*. UTB.

Ziegler, A., & Baker, J. (2013). Talent development as adaptation: The role of educational and learning capital. In S. Phillipson, H. Stoeger, & A. Ziegler (Eds.), *Exceptionality in East Asia: Explorations in the Actiotope Model of Giftedness* (pp. 18–39). Routledge.

Ziegler, A., Chandler, K. L., Vialle, W., & Stoeger, H. (2017). Exogenous and endogenous learning resources in the Actiotope Model of Giftedness and its significance for gifted education. *Journal for the Education of the Gifted, 40*(4), 310–330. https://doi:10.1177/0162353217734376

Ziegler, A., Debatin, T., & Stoeger, H. (2019). Learning resources and talent development from a systemic point of view. *Ann N Y Acad Sci, 1445*(1), 39–51. https://doi.org/10.1111/nyas.14018

Ziegler, A., Dresel, M., Schober, B., & Stoeger, H. (2005). Ulm Motivational Test Battery (UMTB): Documentation of items and scales. *Ulm Educational Psychological Research Report, 15*.

Ziegler, A., Fidelmann, M., Reutlinger, M., Vialle, W., & Stoeger, H. (2010). Implicit personality theories on the modifiability and stability of the action repertoire as a meaningful framework for individual motivation: A cross-cultural study. *High Ability Studies, 21*, 147–163. https://doi.org/10.1080/135 98139.2010.528924

Ziegler, A., & Phillipson, S. N. (2012). Towards a systemic theory of gifted education. *High Ability Studies, 23*(1), 3–30.

Ziegler, A., & Stoeger, H. (2010). Research on a modified framework of implicit personality theories. *Learning and Individual Differences, 20*(4), 318–326. https://doi.org/10.1016/j.lindif.2010.01.007

Ziegler, A., Vialle, W., & Wimmer, B. (2013). The Actiotope model of giftedness: An introduction to some central theoretical assumptions. In S. Phillipson, H. Stoeger, & A. Ziegler (Eds.), *Exceptionality in East Asia: Explorations in the Actiotope Model of Giftedness* (pp. 1–17). Routledge.

Conflicts of interests

The authors declare that there is not any conflict of interest.

Corresponding author:
Anamaria Mendl, PhD
Department of Educational Psychology and Research on Excellence
University of Erlangen-Nuremberg
Regensburger Straße 160
90478 Nuremberg, Germany
E-mail: anamaria.mendl@fau.de