Development of the Inventory of Supporting for Socio-Emotional Skills, evidence of internal structure controlling for acquiescence

Construção, estrutura interna e controle de aquiescência do Inventário de Suporte às Habilidades Socioemocionais

Felipe **VALENTINI**[1] iD 0000-0002-0198-0958

Leonardo de Barros **MOSE**[1] iD 0000-0002-5328-7442

Isabella de Sousa **RAMOS**[2] iD 0000-0002-3516-7359

Nathalia Martins da **CONCEIÇÃO**[2] iD 0000-0001-8061-7032

Abstract

This study aims to develop the Inventory of Supporting for Socio-Emotional Skills, which is an instrument for assessing the support for socio-emotional development provided by teachers, rated by students. It also aims to gather validity evidence of the internal structure of Inventory of Supporting for Socio-Emotional Skills and to control acquiescence. We elaborated the items, which content were evaluated by six experts. For the internal structure, a sample composed of elementary and high-school students ($N = 443$) was used to answer the instrument. The instrument presented good psychometric parameters after controlling for acquiescence using the Random Intercept Model. The final structure was set as unidimensional. We suggest controlling acquiescence to avoid biased scores. In addition, we make an inventory available that can be helpful for socio-emotional skills development.

Keywords: Assessment; Emotional development; Test bias; Test construction.

Resumo

Este estudo apresenta a construção do Inventário de Suporte às Habilidades Socioemocionais, instrumento criado para o aluno avaliar o suporte ao desenvolvimento socioemocional propiciado pelos docentes. A pesquisa pretendeu apresentar evidências de validade da estrutura interna do Inventário de Suporte às Habilidades Socioemocionais e controlar o viés de aquiescência. Para a construção dos itens, seis juízes avaliaram o conteúdo do Inventário. Para avaliar a estrutura interna, uma amostra composta por estudantes de ensino fundamental e médio (N = 443) respondeu presencialmente ao instrumento. Os resultados demonstraram bons parâmetros psicométricos do Inventário após o controle da aquiescência por meio do Modelo de Interceptos Randômicos. A estrutura foi composta por uma dimensão geral de suporte. Sugere-se o controle da aquiescência com a finalidade de evitar distorções nos escores. Ademais, disponibiliza-se um Inventário que pode ser útil ao desenvolvimento das habilidades socioemocionais dos alunos.

Palavras-chave*: Avaliação; Desenvolvimento emocional; Viés do teste; Construção do teste.*

Socio-emotional skills are psychological characteristics, involved in motivation, self-control and relationship with others, which predict a number of outcomes in adulthood (Goodman, Joshi, Nasim, & Tyler, 2015). They have often been referred to as the abilities people must have to perform the demands of life, school and work (McComas, 2014). The development of these skills is a relevant subject for the promotion and improvement of learning in educational institutions (Santos & Primi, 2014a). School plays a fundamental role in this scenario, as it is the core of the exercise of public education policies, in which the teacher occupies a central place. In addition, there have been a lot of discussions about how socio-emotional skills are associated with better student performance in academic activities (Oberle, Schonert-Reichl, Hertzman, & Zumbo, 2014).

However, the role of educators in supporting socio-emotional skills is still a topic little explored in the literature. How can a teacher's empathy, for example, impact students' interpersonal skills? Answering this question still requires an earlier step: how to measure the supportive characteristics of teachers? This investigation focuses on this last question and seeks to build a support scale for the development of socio-emotional skills, so that it can assist in the evaluation of the support of teachers.

Socio-emotional skills and development support

Socio-emotional skills are a set of mainly non-cognitive skills that impact on student performance at school, either in educational institutions, the family environment or their social setting (Primi, Santos, John, & De Fruyt, 2016). Emotional regulation and perseverance are examples of this kind of ability, which contributes – regardless of any change in cognitive skills –, in the student's focus on school activities (Santos & Primi, 2014b). It is noteworthy that socio-emotional skills are individual characteristics present throughout development. However, in this investigation, we will focus on such skills in the school setting, therefore, students' socio-emotional skills and teachers' support.

Even though there is an already known relationship between intelligence and some outcomes in academic life (Valentini & Laros, 2014), there is a remain part of the performance that cannot be explained by cognitive factors. As a result, socio-emotional skills are psychological characteristics that impact on learning and performing tasks in a more satisfactory way. The literature reports evidence of the effect of socio-emotional skills, simultaneously with cognitive skills, on academic achievement and on the reduction of school dropout rates (Duckworth & Seligman, 2005; Santos & Primi, 2014b). Thus, the concept of socio-emotional ability is associated with individual psychological characteristics that lead the individual to perform better, learn more or be more successful in some activity.

F. VALENTINI et al.

2

Estud. psicol. I Campinas I 37 I e180161   2020

The models for the socio-emotional skills mainly involve personality characteristics. The project developed by Collaborative for Academic, Social, and Emotional Learning (Collaborative for Academic Social and Emotional Learning, 2015) proposes a model with five central variables: self-awareness, self-management, social awareness, relationship skills and responsible decision making.

The project developed by the Ayrton Senna Institute proposes a five-dimensional structure similar to the Big Five (Primi et al., 2016; Santos & Primi, 2014a). Although not consensual, the hypothesis that socio-emotional skills can be organized according to this model has been supported by some empirical findings. In the Primi et al. (2016) study the factorial structure of an instrument for assessing socio-emotional skills has been shown to be similar to five major personality dimensions. Furthermore, Pancorbo and Laros (2017) reported convergent validity between an instrument designed to assess socio-emotional skills, SENNA 1.0, and the Reduced Scale of Big Five Personality Factors, a scale for measuring the Big Five dimensions.

The five dimensions of socio-emotional skills that emerge from this taxonomy are: Self-Management, Agreeableness, Engagement with Others, Emotional Resilience, and Openness to the New, and are analogous to the relevant Big Five factors: Conscientiousness, Sociability, Extroversion, Neuroticism, and Openness. Self-management encompasses characteristics of persistence, organization, focus and motivation to complete tasks successfully; Agreeableness involves the skills of empathy, respect and prosocial behaviors; Engagement with Others refers to interpersonal energy, enthusiasm and assertiveness; Emotional Resilience is the ability to coping with stress, tolerance of frustration and self-confidence; and Openness to the New concerns the characteristics of curiosity to learn, imagination and creativity (Primi et al., 2016; Santos & Primi, 2014a).

The five dimensions can also be aggregated into three major axes: interpersonal, intrapersonal and emotional stability. The intrapersonal axis focuses on individual learning and outcome skills and encompasses the dimensions Self-Management and Openness. In the interpersonal axis, the focus is on relationship skills with others and encompasses the dimensions Agreeableness and Engagement with Others. The axis of emotional stability is analogous to the dimension of emotional resilience (De Young, Weisberg, Quilty, & Peterson, 2013).

However, it is well known that not all personality characteristics of these dimensions predict better outcomes in different areas of life. In fact, higher Extroversion scores, for example, correlate negatively with the academic performance (Rosander, Bäckström, & Stenberg, 2011; Valentini & Laros, 2014). In this connection, future studies may still shed light on which socio-emotional skills contribute most to academic performance, as well as their interaction with school variables, such as infrastructure, and teachers skills (technical and didactic).

Another feature of socio-emotional skills is their malleability (Boyce, Wood, & Powdthavee, 2013), especially in childhood. Thus, schools can play a key role in promoting these skills by providing support so that characteristics associated with emotional stability, locus of control, perseverance and tolerance of frustration be developed through public education policies. However, studies in this area are still relatively new and we have not found any specific instrument to assess the support of school and teachers in the development of socio-emotional skills.

Moreover, specifically about the support of school and teachers for the development of students' socio-emotional skills, we are not aware of a specific model for this purpose. However, there are some models more related to the general climate of the institution (Vinha et al., 2016) or teacher styles associated with the technical aspects of teaching (Marchezini-Cunha & Tourinho, 2010). Vinha et al. (2016) model proposes a series of variables associated with the way principals, teachers, staff and students deal with everyday out-of-class situations, such as conflict management and rulemaking. In this connection, the

3

variables listed by the Vinha et al. (2016) model are very specific and not generalizable to support socio-emotional development.

The Marchezini-Cunha & Tourinho (2010) model is generalist, classifying the teaching style in three dimensions: aggressive, in which the teacher only seeks to implement his rights; passive style related to the implementation only of the rights of others; and the assertive style that combines both. However, this model is specific to the didactic style and it is focused on formal education and curricular subjects. In addition, the central characteristics of the three dimensions can be understood from a more general model of parenting styles (Darling & Steinberg, 1993; Morris, Cui, & Steinberg, 2013).

In the styles model (Darling & Steinberg, 1993; Morris, Cui, & Steinberg, 2013), styles are understood from two orthogonal axes: demand, which encompasses characteristics of demanding attainment of goals and objectives and clear establishment of desirable behaviors; and the support axis associated with the emotional, empathy and flexibility aspects. The interaction of these two axes establishes four styles: authoritative (high demand and high support), authoritarian (high demand and low support), negligent (low demand and low support), and permissive (low demand and high support).

It may be seen that the styles model (Darling & Steinberg, 1993) may, theoretically, encompass the model of Marchezini-Cunha and Tourinho (2010). The Aggressive dimension can be understood as low support and high demand (within the authoritarian style); the Passive dimension is characterized by high support and low demand (as well as the permissive style); and the Assertive dimension involves high support and high demand (analogous to the authoritative style). As it is more generic and is supported by more international studies, the style model (Darling & Steinberg, 1993) will be used in this investigation as a basis for the development of the instrument to support social and emotional development.

However, it is emphasized that permissive and negligent styles are difficult to differentiate in practice, probably due to the negative burden they carry: people who evaluate their supervisors, teachers or parents as permissive or negligent will usually do it on account of negativity without sharply discriminating the two styles. Therefore, in the present investigation, we will not distinguish between these two styles in the construction of the items.

Therefore, this investigation aims to develop of an inventory of teacher's support for the student's socio-emotional development, as well as investigate evidence of content validity and internal structure. It is noteworthy that the construction of this scale involves the interaction between five dimensions of socio-emotional skills (Self-Management, Agreeableness, Engagement with Others, Emotional Resilience and Openness to the New) and three styles (Supporter, Demanding, Negligent/Permissive). As hypotheses, we established that: Hypothesis 1a: The factorial structure of the instrument will consist of three dimensions, depending on the styles of demand, support and negligence/permissiveness (Darling & Steinberg, 1993; Morris et al., 2013). Hypothesis 1b: The instrument will consist of five dimensions according to the five domains of socio-emotional skills. (Primi et al., 2016). Hypothesis 1c: The instrument's factorial structure will be organized into three dimensions, one for each axis of socio-emotional skills: interpersonal, intrapersonal and emotional stability (De Young et al., 2013).

Self-reporting and hetero-reporting scales are susceptible to several response biases, which concern to the tendency to endorse items beyond its content, among which, acquiescence (or wording bias) refers to a tendency to positively confirm all items (even inverted items), despite their content (Rammstedt & Farmer, 2013). For example, for an item "I am careful", one will endorse the highest category on the scale (5, "characterizes me very much"), then he endorses the category four ("it characterizes me") for the item "I am negligent". Negligent and careful adjectives have antagonistic meaning and, on the scale, opposite responses would be expected (e.g., 1 and 5 or 2 and 4, or 3 and 3). However, positive endorsements for

F. VALENTINI et al.

4

both items, in the example, point out to a tendency for indiscriminate and idiosyncratic agreement, called acquiescence. Such bias can seriously contaminate the internal structure of the scale and the correlations with external variables (Valentini, 2017). One of the most common problems in this context is the overestimation of a general dimension, reducing the discriminant validity of factors. For the estimation and control of this bias, it is usually sought to build items in distinct poles of content (positive and negative keyed), and it is assumed that there is no acquiescence when responses to antagonistic items are diametrically opposed. On the other hand, any deviation from this expectation is somehow estimated as bias.

Thus, the second objective of this investigation is to estimate the acquiescence. The hypotheses established are: Hypothesis 2a: Adjustment of models with random intercept acquiescence control will be superior to models without control (Aichholzer, 2014). Hypothesis 2b: Factorial loads of positive items will be overestimated without bias control. After acquiescence control, the loads of these items will be shrunk (Soto, John, Gosling, & Potter, 2008).

To achieve the objectives, this report is divided into two studies: in Study 1, we present the construction of the Inventory of Supporting for Socio-Emotional Skills (ISSE). The instrument is designed in a format in which students evaluate the support received from their teacher. In Study 2 we sought to assess the internal structure of ISSE scores. In addition, we sought to review the control of acquiescence response bias through balanced items in the positive (high support) and negative (low support) poles.

Study 1 – Construction of Socio-Emotional Skills Support Inventory and Content Validity Evidence

Method

The construction of the items was based on the interaction between the socio-emotional ability and the teacher's type of style (demand, support, negligent/permissive). Initially, 80 items were built at the high support (positive) and low support (negative) poles. These items were discussed in two research groups specializing in educational psychology. These groups consist of three teachers, and graduated students in the field of psychometrics and educational psychology. Five sessions of two hours each were held. In addition to adjustments in each of the items, the research groups helped to choose the 60th best items to compose the preliminary version of the instrument. This selection was necessary as some items were very similar.

The 60 items chosen were evaluated by a committee of six experts, who indicated if the item intended to compose the instrument was understandable and relevant. Evaluations were performed using a scale from one to five and analyzed using the Content Validity Coefficient (CVC). This coefficient ranges from 0 to 1, and values below 0.80 may indicate wording problems or theoretical coherence and should at least be reviewed (Hernández-Nieto, 2002). It is noteworthy that the cutoff point is arbitrary and therefore other values are also justifiable.

In the next step, a pilot study was conducted with 70 ninth grade elementary school and first year high school students. Each student answered an objective part with 12 items and a descriptive part with three items. In the objective part, the participants would answer if they understood the item (yes or no) and circle any words they did not understand. In the descriptive part, the student should explain in his own words what he understood about the item. The aim was to evaluate, based on the student's rewriting, any comprehension problems. Thus, each item was objectively evaluated by at least five students; as well as rewritten by at least two students.

5

Results

Out of the 60 items submitted for expert's review, eight items presented CVC below 0.80 for understanding the word, and 13 items presented CVC below the cutoff point for relevance (Hernández-Nieto, 2002). All these items were reviewed by two research groups in three meetings. The changes made sought to increase comprehension and avoid dubious interpretations. For example, the item "My teachers don't care about students" could be interpreted, in the research group's assessment, as referring to teachers who don't care about students, as well as referring to the teacher's literal behavior not to call the student, which would not make theoretical sense[3]. Items which were very long and hard to understand were also changed to shorter versions. For example, the item "My teachers believe it is unnecessary to be tolerant" was changed to "My teachers have no patience with students".

Then, in the pilot study, we realized that negative words were the words that students had the most difficulty understanding. In this connection, the word "discourage" was replaced by "do not stimulate" in two items. Some students seemed to understand as positive keyed an item that was created to evaluate the negative pole of the support. This item, which original version was "My teachers always prefer quiet students" was replaced by "My teachers prevent students from being participative." Table 1 shows some examples of items in their final version and their corresponding dimensions.

Study 2 – ISSE Internal Structure Validity Evidence and Response Bias Control

Method

Participants

Data were collected from a non-probability sample of 464 eighth and ninth grade's and first and second year high school students from ten public and private schools in the State of *Rio de Janeiro*. However, we excluded from the analysis 14 participants who had the same answers for all items (including the negatives items) and seven students with a large number of missing data (more than 20.00% of the questionnaires). Therefore, the final sample consisted of 443 participants, 54.00% female, and age ranging from 12 to 23 years (M = 15.40 years; SD = 1.58). The sample size was defined based on the need to investigate data through structural equations, since more complex models with a larger number of parameters require larger samples to perform this type of analysis (Kline, 2015). In addition, in order to verify if 443 participants were adequate for the study, a Montecarlo simulation was performed on *Mplus* with 60 items, 500 replications and 400 subjects (mimicking the actual conditions of the present investigation). There was no convergence problem in any of the replications. The factor loadings were estimated with low bias (bias less than 1.00%) and the true parameter was contained within the confidence interval of the estimated parameter in at least 93.00% of replications. Therefore, a sample of 400 participants is enough to estimate with high statistical power the parameters of the model with 60 items.

Instruments

Inventory of Supporting for Socio-Emotional (ISSE)

This instrument consisted of the 60 items constructed in Study 1. For each item, the student should assess his/her teacher, using a five-point Likert scale, whose anchorage ranged from "not characteristic" to "fully

F. VALENTINI et al.

---

6 [3] Translator's note: In Portuguese "*não liga*" may mean "does not care" and "does not call on the phone".

characteristic". It is noteworthy that the instrument was composed of positive and negative items, which allows the control of acquiescence response bias. The ISSE aims to evaluate the student's perception of the support they receive from the teacher for their social and emotional development. In this connection, the responses of individuals are directed to a common object. In addition, in the study, all students responded to both versions.

Table 1

Examples of Socio-Emotional Skills Support Inventory items

Dimension / Item	Style	Keyed
Interpersonal Axis		
Agreeableness		
My teachers talk about the importance of understanding each other's feelings	Support	+
My teachers decide the workgroups	Demand	+
My teachers don't care if the students have friends	Authoritarian/negligent	-
Engagement with others		
My teachers keep the students in a good mood	Support	+
My teachers ask students to say their opinions	Demand	+
My teachers prevent students from being participatory	Authoritarian/negligent	-
My teachers make students moody	Authoritarian/negligent	-
Intrapersonal / learning axis		
Opening		
My teachers motivate students to use their imagination	Support	+
My teachers take students to museums (or cultural events)	Demand	+
My teachers are closed to new ideas	Authoritarian/negligent	-
Self-management		
My teachers help students to complete assignments	Support	+
My teachers demand students to be organized	Demand	+
My teachers don't give homework	Authoritarian/negligent	-
Emotional Stability Axis		
Emotional Resilience		
My teachers help the student who is sad	Support	+
My teachers ask students to be patient	Demand	+
My teachers laugh at students	Authoritarian/negligent	-

Procedures

The project was submitted to the Research Ethics Committee of the *Universidade Salgado de Oliveira* and approved as case n° 2.718.503. The instruments were applied in the classroom, to students who voluntarily agreed to their participation. Each application session lasted an average of 40 minutes.

Data analyses were conducted in the context of latent modeling by structural equations modeling. We also sought to control the acquiescence bias through the Random Intercept Model (Maydeu-Olivares & Coffman, 2006). In this model, an additional method factor is estimated for all items, setting all loads at a positive value of 1 (including inverted or negative pole items). Thus, the random intercept is estimated as a latent variable that captures inconsistencies of responses. Consequently we used this Exploratory Structural Equation Modeling (ESEM), in which content factors were estimated in a saturated manner (i.e., with freely estimated cross-factor loads analogous to exploratory factor analysis) and the response bias factor loadings were constrained (i.e., fixed factor loadings as well as fixed covariances in 0 with the content factors, similar to the confirmatory factor

analysis). These procedures aimed to diminish the influence of the response bias on the decision about the structure of the scores, since acquiescence, if not controlled, could add an artificial method factor and make it difficult to recover the actual structure of the data. The content factors were rotated using the Geomin method.

In addition, we calculated the classic indicator of acquiescence (mean of positive and negative items without reversing scores). Values above three (Likert Scale center point) are indicative of acquiescent individuals (i.e., they tend to agree with items regardless of positive or negative content). We tested the correlations between this classic indicator of acquiescence, the random intercept, and the descriptive content dimension.

The classic acquiescence indicator was also used to test the influence on a model with only positive items. With this modeling, we meant to test if the absence of negative items could mitigate the effects of acquiescence. Therefore, we analyzed the data by means of Multiple Indicators, Multiple Causes (MIMIC); and, for model identification, the relationship between the acquiescence indicator and the content factor was set to 0 (as theoretically expected).

In all the analyses, the items were declared as categorical and the parameters estimated by Weighted Least Squares Mean and Variance Adjusted. All the analyses were performed using the software Mplus (8.0 version). To evaluate the fit of the model we used the following indicators and cutoff points (Kline, 2015): chi-square; Tucker-Lewis Index (TLI) > 0.95; Comparative Fit Index (CFI) > 0.95; Root Mean Square Error of Approximation (RMSEA) < 0.05.

Results

To explore the internal structure, three and five dimensional models were tested, according to hypotheses 1a, 1b and 1c. In the first analysis, the three-dimensional model was minimally acceptable (TLI = 0.91 and RMSEA = 0.04); however the resulting factors were not interpretable (i.e. the items did not loaded in coherent sets with the three theoretical support styles, demand and negligence/permissiveness). In addition, some strong factor loads (> 0.4) showed signal at the same pole, even though the contents of the items were from opposite poles. For example, in this model, the items "My teachers ignore student sentiment" and "My teachers keep students good-humored", with content at different poles, both had positive factor loads in the first dimension. For these reasons, hypothesis 1a was not confirmed and the model was discarded.

In the second model, we tested the possibility of structuring the instrument in five dimensions, according to the type of socio-emotional ability (hypothesis 1b). This model was also not plausible, and the dimensions randomly aggregated items of different socio-emotional skills. Thus, hypothesis 1b was not supported either.

In the third model, three dimensions were tested again. To differentiate this strategy from the first model tested, the best items were selected. Therefore, the axes of socio-emotional skills were considered (hypothesis 1c) interpersonal (aggregating kindness and engagement with each other), learning (aggregating self-management and openness to experience) and emotional stability. The following criteria were used: each axis should contain the same number of positive and negative items (to balance the scale), the items should have the largest factor loadings and the greatest diversity of thresholds. Therefore, in subsequent analyses, we chose to use only the 22 best items of the instrument. After the selection of the 22 items, two analyses were performed: one with random intercepts acquiescence control and one without control (Table 2).

The axis model presented high correlations between factors (above 0.95). Even controlling for the acquiescence, the factor correlations remained high and above the average factor loadings (approximately 0.50). These results point to the lack of discriminant validity between factors. Thus, we tried to test a one-factor model with the same 22 items (the adjustment indicators are presented in Table 2, bottom). The TLI and CFI indicators of the one-dimensional models were close to those of the three-dimensional model. Due to parsimony, it was decided for a single dimension solution with positive and negative items.

F. VALENTINI et al.

8

Table 2
Goodness of fit of the three-factor model and the one-dimensional model

Models	χ^2	(gl)	TLI	CFI	RMSEA
Three-factors with 22 selected items					
Random Intercept	492.7	(203)	0.95	0.95	0.06
Without bias control	1190.8	(204)	0.84	0.81	0.10
Unidimensional model					
Random Intercept	507.4	(206)	0.95	0.94	0.06
Without bias control	1192.5	(207)	0.84	0.82	0.10

Note: CFI: Comparative Fit Index; RMSEA: Root Mean Square Error of Approximation; TLI: Tucker-Lewis Index.

In addition to the internal structure, the results in Table 2 clearly point to the need to control response bias through random intercepts, because the fit of this model was much higher than the model without bias control, confirming hypothesis 2a of this study. In addition, the variance of random intercepts was estimated at 0.09. Considering that the factor loadings were set at 1 (as a model assumption), the variance of the intercepts indicates that, on average, approximately 9% of the items variability can be attributed to the response bias, a fact that also underlines the importance of estimating this effect.

In addition, we sought to evaluate the relationships between the random intercept, the classic indicator of Classical Test Theory (CTT) acquiescence, and the general support factor. The factor loadings and correlations of these models are presented in Table 3.

The loadings on the random intercepts are fixed at a common value; so in Table 3 these values are the same for all items except items 15 and 48. For these items, acquiescence does not seem to bear any influence, and their loadings were set at 0 according to the modification indications offered by the software. These items have an extreme content (e.g., item 15, "My teachers don't care if students are bullied"), and perhaps for this reason the response bias is lower (or even null). In this connection, to agree with this item, a subject must evaluate his teacher as "very low support level". In addition, the response will be little influenced by acquiescence.

Regarding the associations between variables, the classic acquiescence indicators and random intercepts showed a correlation close to 1, indicating that this is the same variable. In fact, these are distinct (latent or observed) ways of estimating the same phenomenon. In addition, a correlation close to 0 was expected between teacher's support content factor and acquiescence. These correlations were weak and decreased after estimation of random intercepts, but were not close to 0. Possibly, the random intercept did not remove all the variance of the response bias. Still, these results are satisfactory and indicate the importance of controlling biases through positive and negative items.

We also assessed if acquiescence would bias only positive items. In this case, we sought to mimic a common scenario used in psychology research: the absence of negative items in the instruments. The acquiescence indicator was set as an external variable in the MIMIC model (Table 4).

Table 4 shows that the factor loadings of the model with only positive items and without controlling bias were much higher than those of the complete model (presented in Table 3). That is, there seems to be an overestimation of factor loadings in the model with only positive items. In addition, the loads are significantly reduced after the estimation of the direct regression paths between acquiescence (estimated by CTT) and the items. Such regression effects were significant and, in some cases, exceeded the value of the factor load. This fact points out an important bias to be controlled, although it is an estimated model with only positive items. Therefore, hypothesis 2b was confirmed by analyzing the MIMIC model.

9

Table 3

Factor loadings of models with and without random intercept and correlations with classic acquiescence indicator

Items	Random Intercept		Without Random Intercept
	Random Intercept	General Factor	General Factor
4. My teachers motivate students to use the creativity	0.30	0.63	0.65
5. My teachers demand the best effort from students	0.30	0.53	0.57
12. My teachers do not care about the students	0.30	-0.70	-0.66
14. My teachers say the students can have a wonderful future	0.30	0.54	0.57
15. My teachers do not care if students are bullied	0.00	-0.71	-0.70
18. My teachers demands to always use the same way to get things done	0.30	-0.43	-0.40
23. My teachers motivate students to be optimistic	0.30	0.62	0.65
25. My teachers inspire the students to be confident	0.30	0.73	0.76
28. My teachers believe students are skillful	0.30	0.75	0.77
30. My teachers make students moody	0.30	-0.60	-0.58
33. My teachers ignore the students felling	0.30	-0.66	-0.63
34. My teachers motivate students to use their imagination	0.30	0.51	0.53
35. My teachers ask students to say their opinions	0.30	0.60	0.63
41. My teachers discourages the students	0.30	-0.56	-0.54
43. My teachers believe students are unable to learn	0.30	-0.46	-0.44
45. My teachers motivate students to collaborate with colleges	0.30	0.52	0.55
46. My teachers are not sympathetic with students	0.30	-0.64	-0.62
48. My teachers keep the students in a good mood	0.00	0.62	0.62
50. My teachers yell at students	0.30	-0.45	-0.42
51. My teachers guide students to be polite	0.30	0.58	0.62
53. My teachers laugh at students	0.30	-0.67	-0.64
58. My teachers do not praise the student's creativity	0.30	-0.35	-0.32
Correlations			
	FG	FInterRand	FG
FG			
FInterRand	0 (fixo)		-
Aquiec_TCT	0,22	0,99	0,31

Note: FG: content general factor; FInterRand: random intercept factor for acquiescence control; Aquiec_CTT: classic acquiescence indicator, calculated by the average of the positive and negative items (without reversing the score).

Table 4

Multiple Indicators, Multiple Causes (MIMIC) model for controlling positive items acquiescence

Items	Without bias control	MIMIC for controlling bias	
		General factor	Acq_TCT
item 14	0.60	0.57	0.26
item 23	0.69	0.52	0.45
item 25	0.80	0.64	0.47
item 28	0.77	0.75	0.31
item 34	0.58	0.39	0.44
item 35	0.68	0.48	0.49
item 4	0.69	0.52	0.46
item 45	0.62	0.40	0.48
item 48	0.56	0.46	0.33
item 5	0.64	0.48	0.43
item 51	0.67	0.48	0.47

Note: Acq_CTT: classic acquiescence indicator, calculated by the average of the positive and negative items (without reversing the score) calculated before selecting positive items. All parameters presented in the table were statistically significant. (p < 0.001).

F. VALENTINI et al.

General discussion

Regarding study 1, the results showed that the 60 items of the first version of the instrument presented adequate indicators of content validity. The rigorous method of writing items and the evaluation of their quality should be highlighted. In this process, the items were developed based on five socio-emotional skills according to the Big Five personality model (Primi et al., 2016; Santos & Primi, 2014a) and styles of support, demand and neglect (Darling & Steinberg, 1993; Morris et al., 2013). Such perspective allowed the construction of a set of items of different content.

The evaluation of the items by three independent groups (a committee of two research groups; six experts, plus a pilot study with 70 students) allowed to focus on the main instrument limitation: the negative items. On the one hand, some psychometrists suggest the elaboration of positive keyed items (Irwing & Hughes, 2018; Suárez-Alvarez et al., 2018), due to the difficulty in answering inverted items. In fact, in our sample, students in this first pilot study showed greater difficulties in understanding negative items, especially when written with prefix (for example, "discourage" and "dishearten" rather than the negative form "do not encourage" and "do not cheer up"). On the other hand, negative items allow response bias control (Maydeu-Olivares & Coffman, 2006; Rammstedt & Farmer, 2013). It should be noted that response biases associated with wording bias and acquiescence are also present in instruments with only positive items. In this connection, we chose to keep the negative items, replacing the prefix with the word 'no'.

Regarding dimensions, a multidimensional model was expected (hypotheses 1a, 1b and 1c), which was not confirmed in the analyses. Even controlling for response bias, correlations between factors were high enough to make the discriminant validity of these dimensions unfeasible. Thus, hypothetically, students simply perceive their teachers on a continuum from little to much support, without discriminating the type of support received. Another hypothesis is that teachers do not present, on average, visible differences in the type of support, being empathetic and demanding with similar intensity. Another hypothesis concerns a different uncontrolled response bias in this investigation: the halo effect. The halo effect is an usual bias in hetero-reporting instruments and concerns the tendency to evaluate other people uniformly, without differentiation. (Nisbett & Wilson, 1977). In this case, participants would tend to evaluate teachers equally, forcing a unique supportive dimension. In future research, the use of scale with response bias control strategies such as the use of forced choice items is suggested. Through this type of item, it would be possible to control which dimensions would compose each block, forcing the student to choose items of different dimensions.

The attempt to estimate the response bias in this investigation proved to be important. In all the analyses, a significant part of the variance of the items (> 9%) could be attributed to the response bias. On the final model, without controlling the response bias, the factor loadings were apparently overestimated for the positive keyed items (the positive item loads were estimated on average 4% higher than after controlling the response bias), as well as underestimated for the negative keyed items (on average 5% lower).

The relationship between the CTT bias indicator and the content factor of support decreased after the control of random intercepts (from 0.31 to 0.22). The positive correlation in both cases indicates that more acquiescent participants tend to perceive more support from their teachers. It is noteworthy that, in the model without random intercepts, part of this relationship can be attributed to the unbalance in factor loadings of positive and negative items, which were higher for positive keyed items (mean for positive items = 0.63; mean for negative items = -0.54). Thus, endorsing high categories of positive items weighed more heavily in estimating subjects' scores (since the factorial load is higher) than endorsing low categories of negative items. Controlling the random intercepts, the loads show an improvement in balance (mean for positive items = 0.60; mean for negative items = -0.57) and the relationship between acquiescence and scores

11

Estud. psicol. I Campinas I 37 I e180161 2020

decreases (0.31 to 0.22), removing part of the tendency (likely biased) of more acquiescent participants to have a higher perception of teacher support. MIMIC acquiescence control only on positive items was also important, and, if not controlled, overall dimension loads tend to be even more inflated.

Such results also shed light on the discussion of whether to use negative items in inventories. Suárez-Alvarez et al. (2018) indicated a reduction in factor loadings when using negative items in the instrument and suggested the use of positive items only, but did not control acquiescence by CTT or random intercepts. Our study also showed higher loads when only positive items were used. However, by controlling the response bias the effect is reversed. Therefore, our study supports the literature that points out at the importance of using negative items in scales (Danner, Aichholzer, & Rammstedt, 2015; Primi, Hauck-Filho, Valentini, Santos, & Falk, 2019; Rammstedt & Farmer, 2013; Valentini, 2017). In this connection, perhaps answering Likert item is more complicated than it sounds and it is more cognitive demanding than we think in common sense.

Conclusion

This study aimed to construct an instrument that measures support for socioemotional skills, as well as to control acquiescence response bias through a random intercept model and by MIMIC. After testing different models, including multidimensional models, we obtained an inventory that evaluates support in a single continuum, that is, with a one-dimensional factorial structure. Moreover, after verifying the fit of the model with the acquiescence control and without the control, the relevance of controlling biases in Likert hetero-reporting instruments is indicated.

The limitations of this study are the convenience sampling technique used in data collection, making it impossible to generalize the results. Another limitation refers to instrument items that did not contain opposite semantic pairs. The absence of this type of item may have overestimated the acquiescence bias. In addition, if the instrument had been balanced with opposite pairs of items, bias could have been controlled by simpler CTT-based statistical methods (Primi, Santos, De Fruyt, & John, 2019).

As a future research agenda, we intend to review the inventory items, seeking to emphasize the specific behaviors of the teacher that may be associated with different socio-emotional skills. This strategy can result in multidimensional models that are most useful for psychological and educational assessments. Moreover, in future studies, it is intended to test the use of the forced choice item format. Questionnaires in this format eliminate acquiescence, as the items are set in blocks and the respondent must make comparative judgments between them. Since the participant cannot endorse all the items, the acquiescent bias is eliminated without the need for subsequent analysis to correct the scores (Brown, 2016).

In conclusion, a simple hetero-reporting instrument is available for assessing support for social-emotional skills, with evidence of content validity and internal structure. In addition, because it contains items in both poles, the instrument allows the control of response bias in the estimation of the subjects' scores, which importance of the procedure was evidenced in this investigation.

F. VALENTINI et al.

12

Estud. psicol. I Campinas I 37 I e180161   2020

Contributors

F. VALENTINI was ahead of the research design, data analysis and the manuscript writing. L. B. MOSE contributed on the data analysis and the manuscript writing. I. S. RAMOS and N. M. CONCEIÇÃO contributed on data collection and the manuscript writing.

References

Aichholzer, J. (2014). Random intercept EFA of personality scales. *Journal of Research in Personality*, *53*, 1-4. http://dx.doi.org/10.1016/j.jrp.2014.07.001

Almlund, M., Duckworth, A. L., Heckman, J., & Kautz, T. (2011). Personality psychology and economics. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (pp.1-181, Vol.4). Amsterdam: Elsevier. http://dx.doi.org/10.10 16/B978-0-444-53444-6.00001-8

Boyce, C. J., Wood, A. M., & Powdthavee, N. (2013). Is personality fixed? personality changes as much as "variable" economic factors and more strongly predicts changes to life satisfaction. *Social Indicators Research*, *111*(1), 287-305. http://dx.doi.org/10.1007 /s11205-012-0006-z

Brown, A. (2016). Item response models for forced-choice questionnaires: a common framework. *Psychometrika*, *81*(1), 135-160. http://dx.doi.org/10.1007/s11336-014-9434-9

Collaborative for Academic Social and Emotional Learning. (2015). *Effective social and emotional learning programs*. Chicago: Casel.

Danner, D., Aichholzer, J., & Rammstedt, B. (2015). Acquiescence in personality questionnaires: relevance, domain specificity, and stability. *Journal of Research in Personality*, *57*, 119-130. http://dx.doi.org/10.1016/j.jrp.2015.05.004

Darling, N., & Steinberg, L. (1993). Parenting style as context: as integrative model. *Psychological Bulletin*, *113*, 487-496.

De Young, C. G., Weisberg, Y. J., Quilty, L. C., & Peterson, J. B. (2013). Unifying the aspects of the big five, the interpersonal circumplex, and trait affiliation. *Journal of Personality*, *81*(5), 465-475. http://dx.doi.org/10.1111/jopy.12020

Duckworth, A. L., & Seligman, M. E. P. (2005). Outdoes IQ Self-Discipline in predicting of adolescents. *Psychological Science*, *16*(12), 939-944. http://dx.doi.org/10.1111/j.1467-9280.2005.01641.x

Goodman, A., Joshi, H., Nasim, B., & Tyler, C. (2015). *A review for the early intervention foundation: social and emotional skills in childhood and their long-term effects on adult life*. Londres: Institute of Education.

Hernández-Nieto, R. A. (2002). *Contributions to statistical analysis*. Mérida: Universidad de Los Andes.

Irwing, P., & Hughes, D. J. (2018). Test Development. In P. Irwing, T. Booth, & D. J. Hughes (Eds.), *The wiley handbook of psychometric testing* (pp.3-48). Hoboken: Wiley.

Kline, R. B. (2015). *Principles and practice of structural equation modeling* (4th ed.). New York: The Guilford Press.

Marchezini-Cunha, V., & Tourinho, E. Z. (2010). Assertividade e autocontrole: interpretação analítico-comportamental. *Psicologia*: *Teoria e Pesquisa*, *26*(2), 295-304.

Maydeu-Olivares, A., & Coffman, D. L. (2006). Random intercept item factor analysis. *Psychological Methods*, *11*, 344-362. http://dx.doi.org/10.1037/1082-989X.11.4.344

McComas, W. F. (2014). 21st-Century Skills. In W. F. McComas (Ed.), *The language of science education*. Rotterdam: SensePublishers. http://dx.doi.org/10.1007/978-94-6209-497-0_1

Morris, A. S., Cui, L., & Steinberg, L. (2013). Parenting research and themes: what we have learned and where to go next. In R. E. Larzelere, A. S. Morris, & W. A. Harrist (Eds.), *Authoritative parenting: synthesizing nurturance and discipline for optimal child development.* (pp.35-58). Washington: American Psychological Association. http://dx.doi.org/10.1037/13948-003

Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology*, *35*, 250-256.

Oberle, E., Schonert-Reichl, K. A., Hertzman, C., & Zumbo, B. D. (2014). Social-emotional competencies make the grade: predicting academic success in early adolescence. *Journal of Applied Developmental Psychology*, *35*(3), 138-147. http://dx.doi.org/10.1016/ j.appdev.2014.02.004

Pancorbo, G., & Laros, J. A. (2017). Validity evidence of the Social and Emotional Nationwide Assessment (SENNA 1.0) inventory. *Paideia*, *27*(68), 339-347. http://dx.doi.org/10.1590 /1982-43272768201712

13

Primi R., Hauck-Filho N., Valentini F., Santos D., & Falk C. F. (2019) Controlling acquiescence bias with multidimensional IRT modeling. In M. Wiberg, S. Culpepper, R. Janssen, J. González, & D. Molenaar (Eds.), *Quantitative Psychology*: *IMPS 2018* (pp.39-52). Cham: Springer.

Primi, R., Santos, D., De Fruyt, F., & John, O. P. (2019). Comparison of classical and modern methods for measuring and correcting for acquiescence. *British Journal of Mathematical and Statistical Psychology*, 1-19. http://dx.doi.org/10.1111/bmsp.12168

Primi, R., Santos, D., John, O. P., & De Fruyt, F. (2016). Development of an inventory assessing social and emotional skills in Brazilian youth. *European Journal of Psychological Assessment*, *32*(1), 5-16. http://dx.doi.org/10.1027/1015-5759/a000343

Rammstedt, B., & Farmer, R. F. (2013). The impact of acquiescence on the evaluation of personality structure. *Psychological Assessment*, *25*(4), 1137-1145. http://dx.doi.org/ 10.1037/a0033323

Rosander, P., Bäckström, M., & Stenberg, G. (2011). Personality traits and general intelligence as predictors of academic performance: a structural equation modelling approach. *Learning and Individual Differences*, *21*(5), 590-596. http://dx.doi.org/10.1 016/j.lindif.2011.04.004

Santos, D., & Primi, R. (2014a). *Desenvolvimento socioemocional e aprendizado escolar: uma proposta de mensuração para apoiar políticas públicas*. São Paulo: Instituto Ayrton Senna.

Santos, D., & Primi, R. (2014b). *Preliminary results of the social and emotional skills measurement Project in Rio de Janeiro*. São Paulo: Instituto Ayrton Senna.

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2008). The developmental psychometrics of big five self-reports: acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology*, *94*(4), 718-737. http://dx.doi.org/10.1037/0022-3514.94.4.718

Suárez-Alvarez, J., Pedrosa, I., Lozano, L. M., García-Cueto, E., Cuesta, M., & Muñiz, J. (2018). Using reversed items in likert scales: a questionable practice. *Psicothema*, *30*(2), 149-158. http://dx.doi.org/10.7334/psicothema2018.33

Valentini, F. (2017). Influência e controle da aquiescência na análise fatorial: editorial. *Avaliação Psicológica*, *16*(2), 120-123. http://dx.doi.org/10.15689/ap.2017.1602.ed

Valentini, F., & Laros, J. A. (2014). Inteligência e desempenho acadêmico: revisão da literatura. *Temas em Psicologia*, *22*(2), 285-299. http://dx.doi.org/10.9788/TP2014.2-02

Vinha, T. P., Morais, A., Tognetta, L. R. P., Azzi, R. G., Aragão, A. M. F., Marques, C., A. E., … Bozza, T. C. L. (2016). O clima escolar e a convivência respeitosa nas instituições educativas. *Estudos em Avaliação Educacional*, *27*(64), 96-127. http://dx.doi.org/10.18 222/eae.v27i64.3747

F. VALENTINI et al.

14