Original Article

A New Measure for the Assessment of Student–Teacher Connectedness in Adolescence

Irene García-Moya[1] , Fiona Brooks[2], and Carmen Moreno[1]

[1]Departamento de Psicología Evolutiva y de la Educación, Universidad de Sevilla, Spain
[2]Faculty of Health, University of Technology Sydney, Australia

Abstract: Student–teacher relationships have received a lot of attention in the social sciences. It is well-known that relationships with teachers tend to become more distant during adolescence, but most adolescents single out specific teachers they feel connected to. To better capture this latter phenomenon, the aim of this paper is to propose and evaluate the psychometric properties of a new measure for the assessment of student–teacher connectedness in adolescent samples from two countries. Samples consisted of 2,323 adolescents from England and 8,000 adolescents from Spain, who had answered a specific package of questions for the assessment of student–teacher connectedness included in the 17/18 edition of the WHO-collaborative survey Health Behaviour in School-aged Children (HBSC) in the aforementioned two countries. After analyses of item performance and factorial structure, we obtained a 12-item scale with high internal consistency and a factorial structure consistent with the definition of student–teacher connectedness used for scale development in both datasets. Some evidence of validity was also obtained (the scale correlated positively with teacher support, school satisfaction and life satisfaction and negatively with school-related stress). Finally, invariance analysis (configural, metric, and scalar invariance) supported an equivalent functioning of the scale in England and Spain.

Keywords: student–teacher connectedness, adolescence, scale development, invariance

Student–teacher relationships are fundamental for students' learning and well-being (McGrath & Van Bergen, 2015; Roorda et al., 2011). It is well-known that student–teacher relationships tend to become more distant during adolescence, but their potential to contribute to students' positive development remains intact (García-Moya, Brooks, et al., 2015; Roorda et al., 2011). In addition, despite the aforementioned general decrease in the perceived quality of relationships with teachers during adolescence, most adolescent students tend to single out specific teachers they feel connected to (Fredricks et al., 2019; Yu et al., 2018).

Connectedness has been linked to "a pervasive drive to form and maintain at least a minimum quantity of lasting, positive and significant relationships" (Baumeister & Leary, 1995, p. 497). Considered to be an important construct to understand relationships at school, school connectedness was defined by Blum and Libbey (2004, p. 231) as "the belief by students that adults in the school care about their learning and about them as individuals." However, there is now consensus that school connectedness is a multidimensional construct, and unpacking broad definitions of school connectedness to separately examine dimensions such as student–teacher connectedness has been recommended (García-Moya, Bunn, et al., 2019).

General assessments of the quality of relationships with teachers are abundant in the literature. However, student–teacher connectedness requires a distinct level of analysis. Instead of focusing on general feelings about relationships with teachers, an instrument on student–teacher connectedness should capture the availability of a significant relationship with a teacher the student feels connected to and assess its affective quality. An instrument that can provide this complementary information is of particular interest in adolescence, when the transition from primary to secondary education means that students forge relationships that can differ in quality with a considerable number of teachers (typically, one specialist teacher for each subject).

The Middle Years Development Instrument (MDI) is an example of an interesting instrument using an appropriate level of analysis in its subscale for the assessment of connectedness with adults at schools, which focuses on the availability of at least one meaningful connection (Schonert-Reichl et al., 2013). However, this scale focuses on connectedness with adults (teachers or other adults at school), and therefore it does not allow for a specific assessment of student–teacher connectedness.

Among the measures of school connectedness identified in a recent review (García-Moya, Bunn, et al., 2019), two

instruments included a specific subscale for the assessment of connectedness with teachers: the Hemingway Measure of Adolescent Connectedness (Karcher & Lee, 2002) and the Maryland Safe and Supportive Schools Climate Survey (Bradshaw et al., 2014). However, only the Hemingway Measure of Adolescent Connectedness phrased items at the individual level needed to capture student–teacher connectedness, that is, with a focus on each student's individual perception of their relationship with teachers. Although the Hemingway Measure of Adolescent Connectedness allows for reliable and valid assessments of connectedness referred to different domains of life, rather than focusing on the perceived affective quality of the relationship, its scale on teacher connectedness stresses a different aspect: to what extent students value and make efforts to develop a positive relationship with teachers.

Therefore, there is a gap when it comes to instruments specifically developed to assess student–teacher connectedness as the availability of at least one teacher the student connects to and the perceived quality of such connections. The present study presents a measure for the assessment of student–teacher connectedness in adolescence that can contribute to filling this gap.

The Three Dimensions of Student–Teacher Connectedness

We wanted our measure of student–teacher connectedness to reflect essential attributes of connectedness with teachers from the adolescent students' perspective that may be relevant across different cultural and educational contexts. For that reason, the scale development process was supported by a cross-national qualitative study, for which students from England and Spain were selected. Findings from this study (García-Moya, Brooks, et al., 2020) suggest that a humanizing type of relationships is fundamental for student–teacher connectedness, and more specifically that three main attributes can be used to describe such humanizing type of relationship both in England and Spain: (1) individualized personal interactions; (2) empathy and perspective taking; and (3) support.

The *individualized personal interactions* dimension refers to positive interactions with the teacher in which the human side of the teacher is shown and the student feels they are known and respected as persons by the teacher. Students provided specific examples such as teachers knowing and using their name or preferred name, showing an interest in their hobbies and in the activities they are passionate about (arts, sport, music…) and engaging in informal conversations with them in which they showed an interest in getting to know the students and their opinions. The *empa-*

thy and perspective taking dimension encompasses the student's perception that the teacher is sympathetic and understanding and, more broadly, that the teacher is able to see things from the students' perspective. Teachers being sensitive to students' feelings, and attentive to students' behavior or facial expressions that may indicate that the student is not okay were mentioned as relevant aspects of this dimension. Finally, the *support* dimension captures a perception of the teacher as an approachable person, willing to help with academic and non-academic problems if the student needs it.

Accordingly, the following definition of student–teacher connectedness guided scale development in the present study: "a humanizing kind of personalized teacher–student interaction, including the student's perception that (1) teachers notice and respect them as an individual, (2) are sympathetic and can see things from the student's perspective, and (3) act as supportive figures responsive to the student's needs" (García-Moya, Brooks, et al., 2020, p. 17).

The Present Research

The aim of the present study was twofold: (1) to develop a scale on the availability of and perceived quality of at least one meaningful connection with a teacher based on the three aforementioned dimensions of student–teacher connectedness and (2) to conduct a first pilot study of this new scale in two large samples of adolescent students. With a view to explore the potential for using this scale in cross-cultural studies, a sample from England and a sample from Spain were selected. Specifically, a pool of items was developed and piloted in the context of the 17/18 edition of the WHO collaborative survey Health Behaviour in School-aged Children (HBSC) study in the aforementioned two countries.

Our analyses include an assessment of item performance and factorial structure of the scale using confirmatory factor analysis. In addition, given that cross-cultural differences may affect conceptualization of school connectedness (e.g., Sass et al., 2011), the invariance across countries was also tested.

The main focus of the pilot was to obtain a reliable scale with an equivalent functioning across both countries. Nevertheless, to the extent to which data allowed for it, our report on psychometric properties also includes some preliminary evidence of validity, specifically whether the scale correlates in the expected direction with a general assessment of teacher support and measures of school-related stress, school satisfaction, and life satisfaction (e.g., Danielsen et al., 2009; Jiang et al., 2013; Plenty et al., 2014).

Materials and Methods

Participants

Samples consisted of adolescent students aged 11, 13, and 15 years that had taken part in the pilot study of the Teacher Connectedness Project, whose measures were included in the 2017/2018 edition of the WHO collaborative survey HBSC in England and in Spain. Sample size and basic demographic information for each sample is presented in Table 1.

Measures

A pool of 15 items, 5 items for each of the 3 dimensions of student–teacher connectedness (individualized personal interaction; empathy and perspective taking; and support) were developed, including Spanish and English versions of each item (see Table 2).

Following recommendations for scale development, essential content and format specifications were established prior to item development (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014; Wetzel & Greiff, 2018). Regarding content specifications, item development was guided by the definition of student–teacher connectedness and its dimensions presented in the introduction. In addition, the following aspects were specified: the focus on the individual students' perception must be apparent in all items; all items will be formulated in a direct direction; and absolute terms (always, never, etc.) and negatives will be avoided. As for format specifications, after considering different response formats (yes/no format, 5-point Likert scale and 4-point Likert scale), we decided to ask students to express their degree of agreement with each of the items in a Likert-type scale with 4 answer options, labeled verbally as follows: *strongly disagree, somewhat disagree, somewhat agree*, and *strongly agree*. This decision was made because we wanted the response format to be as simple and unequivocal as possible. Although a yes/no format was appealing for its simplicity, we decided against it because greater differentiation is likely to exist when it comes to students' perceptions of their connectedness with teachers. Items with ordinal responses were deemed to fit the construct of interest best and more likely to provide richer information. We then consider using a 5-point Likert scale, which is a very frequently used format. However, although a middle option meant to indicate that the respondent's perception falls in the middle between agree and disagree theoretically allows for greater differentiation in respondents' answers, respondents have been found to interpret midpoint options in ways different to that intended, including expressing indifference, ambivalence

and avoiding socially undesirable options (Nadler et al., 2015). Because the middle option may detract from an unequivocal interpretation of students' responses, a 4-point Likert scale was finally chosen.

In addition to the proposed student–teacher connectedness items, the HBSC questionnaires included some measures we deemed useful for preliminary validation purposes, mainly: teacher support, school-related stress, school satisfaction, and life satisfaction (for further information about these measures, see Inchley et al., 2018).

Procedure

Anonymous questionnaires were completed online by students during school hours in accordance with the HBSC standardized procedure. Ethical approval was obtained from the relevant committees in England and Spain. Furthermore, our pool of items was assessed by 12 international experts on school and/or adolescent research who rated each item's relevance for the assessment of student–teacher connectedness (*not at all relevant, somewhat relevant*, and *very relevant*).

Statistical analyses were performed using SPSS 25 and Mplus 8.3. For continuous data, recommendations for evaluating item performance include removing items with item-total correlations lower than .50 (DeVellis, 2017). Although this is a useful reference value, for our ordinal data we used a comparative approach and supplemented such criterion with the observation of the polychoric correlation matrix to identify items with comparatively lower item-total and inter-item correlations. Expert ratings guided our examination of item performance, with items rated as "not at all relevant" by any expert being flagged as candidates for exclusion and a final decision being made based on the aforementioned statistical criteria (comparatively low item-total and inter-item correlations) and the remaining experts' ratings.

Table 1. Study sample in England and Spain

	England	Spain
Sample size	2,323	8,000
Distribution by gender		
Boys	47.6%	48.6%
Girls	52.4%	51.4%
Distribution by age group		
11 years	719 (30.95%)	2,391 (29.89%)
13 years	891 (38.36%)	2,827 (35.34%)
15 years	713 (30.69%)	2,782 (34.77%)
Proportion of low SES students	17.6%	17.0%

Note. Low SES refers to a score ≤ 6 in the FAS-III scale (Torsheim et al., 2016).

Table 2. Initial pool of items

Connectedness dimension	Item number	English items	Spanish items
IPI	1	I have at least one teacher that is interested in me as a person	Tengo al menos un profesor/a que se interesa por mí como persona
IPI	2	I have at least one teacher that listens to and respects my opinions or ideas	Tengo al menos un profesor/a que escucha y respeta mis opiniones o ideas
IPI	3	I have at least one teacher that knows my name	Tengo al menos un profesor/a que se sabe mi nombre
IPI	4	I have at least one teacher that pays attention to me as an individual even in a big class	Tengo al menos un profesor/a que me presta atención individual incluso en una clase con muchos alumnos
IPI	5	I have at least one teacher that takes an interest in my hobbies and interests outside of school	Tengo al menos un profesor/a que se interesa por mis aficiones e intereses fuera de la escuela
EPT	6	I have at least one teacher that notices when I struggle to follow the class	Tengo al menos un profesor/a que se da cuenta cuando me cuesta seguir la clase
EPT	7	I have at least one teacher that notices when I'm worried about something	Tengo al menos un profesor/a que nota cuando estoy preocupado/a por algo
EPT	8	I have at least one teacher that tries to see things from my perspective	Tengo al menos un profesor/a que intenta ver las cosas desde mi perspectiva
EPT	9	I have at least one teacher that is sympathetic with me if I feel upset or cry in class	Tengo al menos un profesor/a que es comprensivo/a conmigo si no me encuentro bien emocionalmente o lloro en clase
EPT	10	I have at least one teacher that tries to make me feel better when I'm sad	Tengo al menos un profesor/a que intenta que me sienta mejor cuando estoy triste
S	11	I have at least one teacher that is supportive of me and tries to help me	Tengo al menos un profesor/a que me apoya y trata de ayudarme
S	12	I have at least one teacher that is willing to listen to my problems	Tengo al menos un profesor/a que está dispuesto/a a escuchar mis problemas
S	13	I have at least one teacher that is there for me if I need their help	Tengo al menos un profesor/a que está ahí para mí si necesito su ayuda
S	14	I have at least one teacher that would listen to me if I confided them a problem	Tengo al menos un profesor/a que me escucharía si le contara un problema
S	15	I have at least one teacher that knows how to help me	Tengo al menos un profesor/a que sabe cómo ayudarme

Note. IPI = Individualized personal interaction; EPT = Empathy and perspective taking; S = Support.

Once item functioning had been examined, we run confirmatory factor analysis (CFA) using Mplus 8.3. As recommended to obtain a more robust test of the model in CFA using structural equation modeling (SEM), a competing model estimation strategy was used (Hair et al., 2009), in which we compared a 1-factor solution and a 3-correlated-factor solution based on the theoretical assumptions presented in the introduction. To take into consideration the ordinal nature of the items, we employed categorical confirmatory factor analysis using the weighted least squares means and variance adjusted (WLSMV) estimator, so that analysis is based on a polychoric correlation matrix with adjusted means and variance estimates. Goodness of fit was evaluated using chi-square differences ($\Delta\chi^2$), which were calculated using the difftest option in Mplus, and the approximate goodness-of-fit indices: Tucker-Lewis Index (TLI; also known as Nonnormed Fit Index, NNFI), Comparative Fit Index (CFI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR). Values of TLI/NNFI and CFI higher than .90 (McDonald & Ho, 2002) and RMSEA < .08 (Browne & Cudeck, 1993) have been considered to be indicative of acceptable fit, while Hu and Bentler (1999) proposed more stringent criteria for good fit: values of .95 or higher for CFI

and NNFI and RMSEA values lower than .06. SRMR of .08 or less is also desired in a well-fitting model (Hu & Bentler, 1999).

Several authors have warned that the aforementioned reference values for approximate fit indices were developed in studies using continuous variables, with some caution needed when applying them to ordinal data (e.g., Kline, 2016; Shi et al., 2020). For example, a recent study concluded that RMSEA tends to increasingly reject the close fit hypothesis in large samples as the number of categories increases, and recommended SRMR as a preferable option to assess the degree of misfit in ordinal factor analysis models (Shi et al., 2020). We also took into consideration the following recommendations for assessing goodness of fit in analyses with ordinal Likert type variables: not to rely on a single index (Pendergast et al., 2017) and examine correlation residuals (i.e., differences between the sample polychoric correlations and values predicted by the model), for which absolute values < .10 can be interpreted as an indication that the model closely reproduces the sample correlations (Kline, 2016).

Finally, we conducted SEM multi-group analysis to examine the invariance across countries for the model that showed the best goodness of fit in the previous analyses.

Table 3. Item performance summary

Item No	England						Spain					
					Experts						Experts	
	Mdn	IQR	ITC	VR	SR	NR	*Mdn*	IQR	ITC	VR	SR	NR
1	3	2	.766	5	1	0	4	1	.734	5	1	0
2	3	1	.834	5	1	0	4	1	.760	6	0	0
3	4	0	.568	2	3	1	4	0	.513	3	1	2
4	3	1	.756	3	3	0	3	1	.686	5	1	0
5	3	2	.752	2	2	2	3	2	.708	2	4	0
6	3	2	.764	5	0	1	3	1	.699	5	0	1
7	3	1	.791	4	2	0	3	2	.767	6	0	0
8	3	1	.835	3	3	0	3	2	.791	6	0	0
9	3	2	.832	5	1	0	3	1	.799	5	1	0
10	3	2	.858	4	2	0	3	1	.818	5	1	0
11	3	1	.884	4	2	0	3	1	.859	5	1	0
12	3	1	.888	4	2	0	3	1	.840	6	0	0
13	3	1	.898	4	2	0	3	1	.843	5	1	0
14	3	1	.878	5	1	0	3	1	.832	5	1	0
15	3	2	.831	3	3	0	3	1	.822	3	2	1

Note. IQR = interquartile range; ITC = Item-total correlation; VR = Number of experts who rated the item as "Very relevant"; SR = Number of experts who rated the item as "Somewhat relevant"; NR = Number of experts who rated the item as "Not relevant".

For that purpose, we followed the recommendations by Pendergast et al. (2017) on best practices for multi-group invariance confirmatory factor analyses with ordinal Likert-type items in school psychology, which are based on the work by Millsap and Yun-Tein (2004). Specifically, we started by testing configural invariance, to see whether the same number of factors and factor-indicator correspondence provides an adequate fit to both groups (England and Spain). These results are used as a baseline for testing metric invariance, which answers the question of whether factor loadings of each indicator are similar across groups. For that purpose, equality constrains for all factor loadings were imposed simultaneously and the constrained model fit was compared to the previous model. The next step is examining scalar invariance, that with ordinal data involves the examination of item threshold invariance. To test scalar/ threshold invariance, we set intercepts at zero and constrained thresholds to be equal across groups. If thresholds are found to be invariant across groups, the invariance of intercepts can be assumed and scalar invariance is supported.

To evaluate invariance at each successive step, ΔCFI has been considered to be among the most promising approximate fit indices and it is especially recommended in large sample studies since it is less sensitive to sample size and more sensitive to lack of invariance than chi-square and non-redundant with other approximate fit indices (Cheung & Rensvold, 2002; Meade et al., 2008). According to Cheung and Rensvold (2002), a decrease in CFI lower than .01 in the more constrained model and the maintenance of acceptable values for the rest of indices indicate that the

invariance hypothesis can be accepted. Meade et al. (2008) recommend a more stringent cut-off point (change in CFI lower than .002) for metric and scalar invariance. Although this is an area in continuous development where a clear consensus is yet to be achieved, ΔCFI lower than .01 and ΔRMSEA lower than .015 have been considered to be useful criteria when examining invariance in correctly specified models with categorical indicators (Pendergast et al., 2017; Sass et al., 2014), so we used those as reference values for the present study.

Results

Item Performance

Table 3 provides a summary of item performance information, including median, interquartile range (IQR), corrected item-total correlation and expert ratings. In addition, the full correlation matrices, where inter-item correlations can be inspected, can be found in the Appendix (available as Electronic Supplementary Material, see ESM 11). Expert ratings on item relevance flagged items 3, 5, and 6 from the English version and items 3, 6, and 15 from the Spanish version as candidates for exclusion, since at least one expert rated them as "not relevant" (see Table 3).

After the inspection of item performance information, we made the decision to remove items 3 and 6 in both countries. In England, item 3 had the lowest item-scale correlation (.568), comparatively low inter-item correlations (inter-item correlation mean = .556), it was rated as "not

at all relevant" by one expert and the majority of experts did not consider the item was "very relevant." In Spain, item 3 also had the lowest item-scale correlation (.513), comparatively low inter-item correlations (inter-item correlation mean = .486) and it was rated as "not relevant" by two experts. Item 3 also showed low variability in both countries, as indicated by the median and IQR values. Although median and IQR suggested greater variability for item 6, the fact it had also been flagged as "not relevant" by experts in both countries coupled with its item-scale (.764 in England and .699 in Spain) and inter-item correlations (inter-item correlation mean = .684 in England and .621 in Spain) among the lowest in both countries also led us to the decision to remove item 6.

In the case of item 5, the situation was less clear: item 5 was flagged as a potential candidate for exclusion based on experts' ratings in England only and IQR and median suggested sufficient variability, but its item-scale correlation and inter-item correlations were relatively similar to those for item 6 (i.e., among the lowest in both countries). At this point, a tentative decision was made to keep item 5 with a view to further examine its performance in confirmatory factor analysis to make a final decision. Finally, despite having been rated as "not relevant" by an expert in Spain, item 15 was retained because it showed high item-total and inter-item correlations and sufficient variability in both countries.

Confirmatory Factor Analysis

Confirmatory factor analysis was conducted with the remaining 13 items. To allow for a more robust test, we compared the goodness of fit of two competing models based on the theoretical assumptions presented in the introduction: a 1-factor model (representing a global connectedness factor) and an alternative competing model where the three dimensions guiding item development were modeled as three correlated factors (i.e., 3-correlated-factors model). The absolute fit indicator χ^2, the $\Delta\chi^2$, and approximate goodness-of-fit indices for these models are reported in Table 4.

The 1-factor model did not show a good fit to the data in England or Spain. Although most approximate fit indices were consistent with a good fitting model, the RMSEA values (.143 in England and .120 in Spain) were substantially higher than those recommended for an acceptable fit. The $\Delta\chi^2$ between the 3-correlated-factors model and the 1-factor model was consistent with a significant difference in model fit between the two competing nested models. Specifically, although the model's χ^2 value remained significant, which may be explained by the large sample size, comparative fit indices for the 3-correlated-factors model were consistent with a better fitting model: TLI/NNFI

and CFI were above the .95 threshold and higher compared to the 1-factor model, and the decrease in RMSEA values (now closer to the recommended < .08 threshold for acceptable fit) and the lower values of SRMR indicated a reduction of model-data discrepancies compared to the 1-factor model.

Given that caution has been recommended when applying reference values for approximate fit indices in general and for RMSEA in particular to analyses with ordinal data (e.g., Kline, 2016; Shi et al., 2020), additional steps were followed to further assess goodness of fit and item performance in the selected model (the 3-correlated-factors model). Specifically, we run additional analyses to obtain the correlation residual matrix as well as modification indices. Although modification indices can provide some information about model misfit, there is consensus that re-specifying models merely based on modification indices is problematic, since they rely exclusively on empirical criteria; therefore, modification indices must be used in a judicious way, with post hoc modifications being kept to a minimum and only theory-based modifications being considered (Byrne, 2012; Kline, 2016). In this study, the inspection of modification indices was deemed useful to further examine the functioning of item 5, which had been flagged by experts as a candidate for exclusion and for which a tentative decision to retain was made, with a view to make a final decision based on confirmatory factor analysis results. After inspecting modification indices with a focus on item 5, we decided to remove this item from our final model, since results suggested that adding cross-loadings to the empathy factor would be associated with substantial improvements in model fit in both countries.

A graphical representation of the estimated final 12-item model after removing item 5 is provided in Figure 1. Although RMSEA values may still be considered borderline in the case of England, when taking into consideration multiple fit indices, as recommended, the model showed an acceptable fit to the data in both countries, and SRMR, which seems to be a more effective index of model misfit in ordinal factor analysis than RMSEA (Shi et al., 2020), was consistent with a good fitting model (see Table 4). In addition, factor loadings for all items were high, and all correlation residual values were below .10, which is an indication that the model closely reproduced the sample correlations.

Reliability and Preliminary Indications of Validity

Composite reliability for each of the dimensions and for the 12-item scale was calculated. Composite reliability is based on the ratio of explained variance over total variance and it

Table 4. Confirmatory factor analysis: Summary of goodness-of-fit indices

	England			Spain		
	1 factor (13 items)	3-correlated-factors (13 items)	3-correlated-factors (12 items)	1 factor (13 items)	3-correlated-factors (13 items)	3-correlated-factors (12 items)
χ^2	2,071.090	780.764	565.619	6,775.877	3,013.990	1,969.735
p	< .001	< .001	< .001	< .001	< .001	< .001
df	65	62	51	65	62	51
DIFFTEST	$\Delta\chi^2(df)$ = 518.594 (3)	p < .001	n/a†	$\Delta\chi^2(df)$ = 1641.840 (3)	p < .001	n/a†
TLI/NNFI	.975	.990	.993	.970	.986	.990
CFI	.979	.992	.995	.975	.989	.993
RMSEA [90% CI]	.143 [.137, .148]	.087 [.082, .093]	.082 [.076, .088]	.120 [.118, .122]	.081 [.079, .084]	.072 [.070, .075]
SRMR	.030	.017	.014	.030	.019	.014

Note. Due to the adjustments made to χ^2 by the WLSMV estimator, χ^2 values cannot be directly compared between the nested 1-factor and 3-correlated-factor models. $\Delta\chi^2$ to assess changes in model fit between these two competing nested models was calculated using the DIFFTEST function in Mplus 8.3. †The 3 correlated factors (12 items) model was estimated after removal of item 5; because of its different set of observed variables this model is not nested and DIFFTEST is not applicable. $\Delta\chi^2$ = Chi-Square Difference; df = Degrees of Freedom; TLI/NNFI = Tucker-Lewis Index/Nonnormed Fit Index; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; CI = Confidence Interval; SRMR = Standardized Root Mean Square Residual.

is considered to be a better option than Cronbach's α for assessing reliability in confirmatory factor analysis models (for a more detailed description of composite reliability, see Kline, 2016, pp. 313–314). In England, composite reliability based on the WLSMV solution was .91 for individualized personal interactions, .93 for empathy and perspective taking, .97 for support, and .98 for the 12-item scale. In Spain, composite reliability was .88 for individualized personal interactions, .92 for empathy and perspective taking, .96 for support, and .98 for the 12-item scale. It must be noted that when one assesses composite reliability based on WLSMV estimates, the calculation is based on the reliability of the estimated underlying continuous variables. For that reason, we also calculated composite reliability using the procedure recommended by Raykov and Marcoulides for reliability evaluation of categorical items (Raykov & Marcoulides, 2011). In brief, this involved grouping categorical items into parcels and using robust maximum likelihood to benefit from its robustness against deviations from normality. Using this procedure, we obtained the following composite reliability point estimates and 95% confidence intervals for our 12-item scale: .945 (95% CI = .940, .950) in England and .932 (90% CI = .929, .935) in Spain.

To provide some preliminary indications of validity we examined the extent to which the scale functioning was congruent with previous research, specifically whether student–teacher connectedness correlated in the expected directions with other school-related variables and life satisfaction. In England, the scale correlated positively with a general measure of teacher support (r = .50), school satisfaction (r = .35), and life satisfaction (r = .17) and negatively with school-related stress (r = −.13). Similar results were found in Spain. Connectedness scores correlated positively with teacher support (r = .59), school satisfaction (r = .31), and life satisfaction (r = .28) and negatively with

school-related stress (r = −.15). All these correlations were statistically significant (p < .001).

Invariance Across Countries

SEM multi-group analysis with the two countries as groups was used to examine invariance. As shown in Table 5, when the same number of factors and factor-indicator correspondence was modeled in both groups an acceptable fit was obtained, supporting the most basic form of invariance: configural invariance. Both the decrease in CFI (lower than .002) and the maintenance of the remaining indices in acceptable values when imposing cross-group equality constraints on factor loadings is consistent with the metric invariance hypothesis that factor loadings of each indicator are equivalent across the groups. Results are also consistent with scalar invariance, for which we set intercepts at zero and constrained both factor loadings and thresholds to be equal across the groups.

Discussion

The aim of the present study was to develop a scale that can be used in survey research with adolescent students to assess the availability of at least one teacher to which the student feels connected to and the quality of these connectedness experiences based on three main dimensions: individualized personal interactions, empathy and perspective taking and support.

As further explained in the introduction, the proposed measure covers a gap in the literature. In addition, our results indicate that this scale allows measuring student–teacher connectedness reliably in England and
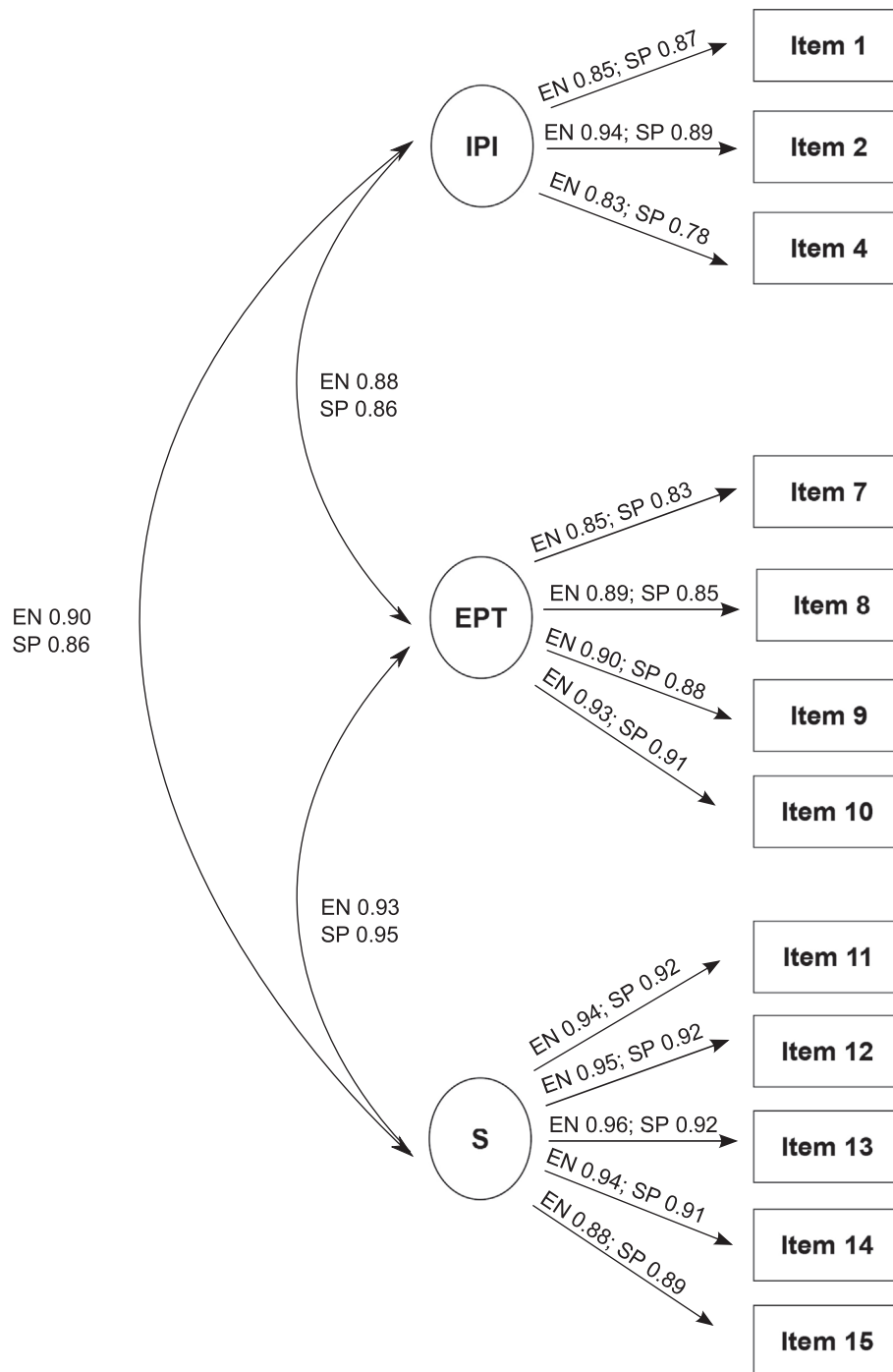
Figure 1. Structural equation model for the student–teacher connectedness scale in England (EN) and Spain (SP) (standardized parameters reported). IPI = Individualized personal interaction; EPT = Empathy and perspective taking; S = Support.

Spain and it reflects a 3-correlated-factors structure, which is consistent with the definition used (García-Moya, Brooks, et al., 2020).

The total score of student–teacher connectedness correlated positively with teacher support, and as expected based on previous research on student–teacher relationships (Danielsen et al., 2009; Jiang et al., 2013), was significantly

and positively associated with school satisfaction and life satisfaction, which provides some preliminary evidence of validity. Student–teacher connectedness also correlated negatively with school-related stress, as also expected (Plenty et al., 2014). Importantly, the correlation with teacher support (.50 in England and .59 in Spain) supports convergent validity but it is not so high as to suggest that the

Table 5. Scale invariance across countries

	Configural invariance	Metric invariance	Scalar/threshold invariance
χ^2	2,572.821	2,529.494	2,796.400
df	102	111	132
CFI	.993	.993	.992
ΔCFI	–	.000	.001
TLI/NNFI	.991	.992	.992
RMSEA [90% CI]	.075 [.072, .077]	.071 [.068, .073]	.068 [.066, .070]
ΔRMSEA	–	.004	.003
SRMR	.014	.014	.017

Note. df = Degrees of Freedom; TLI/NNFI = Tucker-Lewis Index/Nonnormed Fit Index; CFI = Comparative Fit Index; ΔCFI = CFI difference; RMSEA = Root Mean Square Error of Approximation; ΔRMSEA = RMSEA difference; SRMR = Standardized Root Mean Square Residual.

same construct is being measured by both scales. Finally, the reliability of the final scale was high, with composite reliability above .90 in both countries. Reliability was also good for the three dimensions of connectedness. Because the three dimensions correlate significantly with one another, we recommend that the complete scale is used to obtain more comprehensive assessments of student–teacher connectedness.

Finally, invariance analyses suggest that the scale showed an equivalent functioning in adolescents from England and Spain. This is very important given the existing debate on whether cultural differences and educational dissimilarities may affect the meaning of connectedness with teachers in different countries (see e.g., Sass et al., 2011) and suggests that the scale can be used in cross-national studies in the examined countries. Despite the emphasis on the importance of testing measurement invariance before meaningful mean group comparisons and predictive associations with outcomes of interest can be confidently undertaken in cross-national studies (e.g., Brown et al., 2017; Jeong & Lee, 2019), a recent review warns on the limited presence of invariance testing in cross-cultural research (Boer et al., 2018). Therefore, the fact that we found invariance across countries is an important first step for the use of the proposed scale in future cross-cultural studies of student–teacher connectedness that can shed additional light on the similarities and differences between students' experiences of connectedness in England and Spain.

This study has some limitations. Mainly, it is advisable that item selection and validation are performed in independent samples. We used large samples from two different countries in the present study but unfortunately we were not able to collect data from two independent samples in each country. We also used the competing modeling strategy for confirmatory factor analyses to minimize risks associated with the use of confirmatory techniques, which contributes to the robustness of the results, and followed several steps recommended for increasing validity early in the scale development process (Gehlbach & Brinkworth,

2011). Specifically, our construct definition was informed by a scoping review on the conceptualization of school and teacher connectedness as well as by a cross-cultural qualitative study capturing main attributes of this construct. Furthermore, the initial pool of items was assessed by 12 international experts on school and/or adolescent research whose ratings supported content validity and have guided item selection.

Nevertheless, new validation studies in independent samples and additional research to explore the possibility to develop a shorter version of this student–teacher connectedness scale that can be used in large-scale surveys would be worth pursuing. Most of the items in our original item pool were retained, resulting in a 12-item scale. In our opinion, given the high inter-item correlations and composite reliability for the scale and its dimensions found in the present study, there seems to be room to test a shorter version with a reduced number of items per factor as a next step to continue with the validation and refinement of this scale in future studies. Furthermore, the evidence of invariance across countries provided in the present study can be complemented by testing the scale in additional countries. Finally, although we relied on qualitative work aimed at identifying common central aspects of connectedness with teachers in adolescent boys and girls of different age groups, it would be advisable to empirically examine the assumption of equivalent functioning across gender and different age groups by means of additional invariance studies.

Electronic Supplementary Materials

The electronic supplementary material is available with the online version of the article at https://doi.org/10.1027/1015-5759/a000621

ESM 1. This file contains the Mplus syntax input and output of the item performance analyses performed in the England dataset.

ESM 2. This file contains the Mplus syntax input and output of the item performance analyses performed in the Spain dataset.

ESM 3. This file contains the Mplus syntax input and output of the confirmatory factor analyses performed in the England dataset. These are presented in the following order: 3-correlated-factors model, 1-factor model and calculation of difftest, 3-correlated-factors model including correlation residuals and Modification Indices, and final 12-item 3-correlated-factors model.

ESM 4. This file contains the Mplus syntax input and output of the confirmatory factor analyses performed in the Spain dataset. These are presented in the following order: 3-correlated-factors model, 1-factor model and calculation of difftest, 3-correlated-factors model including correlation residuals and Modification Indices, and final 12-item 3-correlated-factors model.

ESM 5. This file contains the Mplus syntax input and output of the additional analyses for calculating composite reliability in England following Raykov and Marcoulides' procedure (2011).

ESM 6. This file contains the Mplus syntax input and output of the additional analyses for calculating composite reliability in Spain following Raykov and Marcoulides' procedure (2011).

ESM 7. This file contains the SPSS syntax and output correlation tables for the 12-item teacher connectedness scale, teacher support, school satisfaction, life satisfaction, and school pressure in England and Spain.

ESM 8. This file contains the Mplus syntax input and output of the multi-group analysis of configural invariance.

ESM 9. This file contains the Mplus syntax input and output of the multi-group analysis of metric invariance.

ESM 10. This file contains the Mplus syntax input and output of the multi-group analysis of scalar/threshold invariance.

ESM 11. Appendix. Correlations matrices for the 15 teacher connectedness items in the England and Spain datasets.

References

American Educational Research Association.American Psychological Association.National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychology Bulletin, 117*(3), 497–529. https://doi.org/10.1037/0033-2909.117.3.497

Blum, R. W., & Libbey, H. P. (2004). Executive summary. *Journal of School Health, 74*(7), 231–232. https://doi.org/10.1111/j.1746-1561.2004.tb08278.x

Boer, D., Hanke, K., & He, J. (2018). On detecting systematic measurement error in cross-cultural research: A review and critical reflection on equivalence and invariance tests. *Journal of Cross-Cultural Psychology, 49*(5), 713–734. https://doi.org/10.1177/0022022117749042

Bradshaw, C. P., Wassdorp, T. E., Debnam, K. J., & Johnson, S. L. (2014). Measuring school climate in high schools: A focus on safety, engagement and the environment. *Journal of School Health, 84*(9), 593–604. https://doi.org/10.1111/josh.12186

Brown, G. T. L., Harris, L. R., O'Quin, C., & Lane, K. E. (2017). Using multi-group confirmatory factor analysis to evaluate crosscultural research: Identifying and understanding non-invariance. *International Journal of Research & Method in Education, 40*(1), 66–90. https://doi.org/10.1080/1743727X.2015.1070823

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Sage.

Byrne, B. M. (2012). *Structural equation modeling with Mplus. Basic concepts, applications, and programming*. Routledge.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Danielsen, A. G., Samdal, O., Hetland, J., & Wold, B. (2009). School-related social support and students' perceived life satisfaction. *The Journal of Educational Research, 102*(4), 303–320. https://doi.org/10.3200/JOER.102.4.303-320

DeVellis, R. F. (2017). *Scale development: Theory and applications*. Sage.

Fredricks, J. A., Parr, A. K., Amemiya, J. L., Wang, M.-T., & Brauer, S. (2019). What matters for urban adolescents' engagement and disengagement in school: A mixed-methods study. *Journal of Adolescent Research, 34*(5), 491–527. https://doi.org/10.1177/0743558419830638

García-Moya, I., Brooks, F., & Moreno, C. (2020). Humanizing and conducive to learning. An adolescent students' perspective on the central attributes of positive relationships with teachers. *European Journal of Psychology of Education, 35*(1), 1–20. https://doi.org/10.1007/s10212-019-00413-z

García-Moya, I., Brooks, F., Morgan, A., & Moreno, C. (2015). Subjective well-being in adolescence and teacher connectedness. A health asset analysis. *Health Education Journal, 74*(6), 641–654. https://doi.org/10.1177/0017896914555039

García-Moya, I., Bunn, F., Jiménez-Iglesias, A., Paniagua, C., & Brooks, F. (2019). The conceptualization of school and teacher connectedness in adolescent research: A scoping review of literature. *Educational Review, 71*(4), 423–444. https://doi.org/10.1080/00131911.2018.1424117

Gehlbach, H., & Brinkworth, M. E. (2011). Measure twice, cut down error: A process for enhancing the validity of survey scales. *Review of General Psychology, 15*(4), 380–387. https://doi.org/10.1037/a0025704

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (2009). *Análisis multivariante*. Pearson, Prentice-Hall.

Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Inchley, J., Currie, D., Cosma, A., & Samdal, O. (2018). *Health Behaviour in School-aged Children (HBSC) study protocol: Background, methodology and mandatory items for the 2017/18 survey*. CAHRU.

Jeong, S., & Lee, Y. (2019). Consequences of not conducting measurement invariance tests in cross-cultural studies: A review of current research practices and recommendations. *Advances in Developing Human Resources, 21*(4), 466–483. https://doi.org/10.1177/1523422319870726

Jiang, X., Huebner, S. E., & Siddall, J. (2013). A short-term longitudinal study of differential sources of school-related

social support and adolescents' school satisfaction. *Social Indicators Research, 114*(3), 1073–1086. https://doi.org/10.1007/s11205-012-0190-x

Karcher, M., & Lee, Y. (2002). Connectedness among Taiwanese middle school students: A validation study of the Hemingway Measure of Adolescent Connectedness. *Asia Pacific Education Review, 3*(1), 92–114. https://doi.org/10.1007/BF03024924

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.

McDonald, R. P., & Ho, M. R. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods, 7*(1), 64–82. https://doi.org/10.1037/1082-989X.7.1.64

McGrath, K. F., & Van Bergen, P. (2015). Who, when, why and to what end? Students at risk of negative student–teacher relationships and their outcomes. *Educational Research Review, 14*, 1–17. https://doi.org/10.1016/j.edurev.2014.12.001

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology, 93*(3), 568–592. https://doi.org/10.1037/0021-9010.93.3.568

Millsap, R. E., & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research, 39*(3), 479–515. https://doi.org/10.1207/S15327906MBR3903_4

Nadler, J. T., Weston, R., & Voyles, E. C. (2015). Stuck in the middle: The use and interpretation of mid-points in items on questionnaires. *The Journal of General Psychology, 142*(2), 71–89. https://doi.org/10.1080/00221309.2014.994590

Pendergast, L. L., von der Embse, N., Kilgus, S. P., & Eklund, K. R. (2017). Measurement equivalence: A non-technical primer on categorical multi-group confirmatory factor analysis in school psychology. *Journal of School Psychology, 60*, 65–82. https://doi.org/10.1016/j.jsp.2016.11.002

Plenty, S., Östberg, V., Almquist, Y. B., Augustine, L., & Modin, B. (2014). Psychosocial working conditions: An analysis of emotional symptoms and conduct problems amongst adolescent students. *Journal of Adolescence, 37*(4), 407–417. https://doi.org/10.1016/j.adolescence.2014.03.008

Raykov, T., & Marcoulides, G. A. (2011). Procedures for estimating reliability. In T. Raykov & G. A. Marcoulides (Eds.), *Introduction to psychometric theory* (pp. 147–182). Taylor & Francis.

Roorda, D. L., Koomen, H. L., Spilt, J. L., & Oort, F. J. (2011). The influence of affective teacher-student relationships on students' school engagement and achievement: A meta-analytic approach. *Review of Educational Research, 81*(4), 493–529. https://doi.org/10.3102/0034654311421793

Sass, D. A., Castro-Villarreal, F., McWhirter, B. T., McWhirter, E. H., & Karcher, M. J. (2011). A cross-cultural assessment of school connectedness: Testing measurement invariance with United States and Chilean adolescents. *Journal of Psychoeducational Assessment, 29*(5), 428–442. https://doi.org/10.1177/0734282910391067

Sass, D. A., Schmitt, T. A., & Marsh, H. W. (2014). Evaluating model fit with ordered categorical data within a measurement invariance framework: A comparison of estimators. *Structural Equation Modeling: A Multidisciplinary Journal, 21*(2), 167–180. https://doi.org/10.1080/10705511.2014.882658

Schonert-Reichl, K. A., Guhn, M., Gadermann, A. M., Hymel, S., Sweiss, L., & Hertzman, C. (2013). Development and validation of the Middle Years Development Instrument (MDI): Assessing children's well-being and assets across multiple contexts. *Social Indicators Research, 114*(2), 345–369. https://doi.org/10.1007/s11205-012-0149-y

Shi, D., Maydeu-Olivares, A., & Rosseel, Y. (2020). Assessing fit in ordinal factor analysis models: SRMR vs. RMSEA. *Structural Equation Modeling: A Multidisciplinary Journal, 27*(1), 1–15. https://doi.org/10.1080/10705511.2019.1611434

Torsheim, T., Cavallo, F., Levin, K. A., Schnohr, C., Mazur, J., Niclasen, B., & Currie, C., The FAS Development Study Group. (2016). Psychometric validation of the revised Family Affluence Scale: A latent variable approach. *Child Indicators Research, 9*, 771–784. https://doi.org/10.1007/s12187-015-9339-x

Wetzel, E., & Greiff, S. (2018). The world beyond rating scales. Why we should think more carefully about the response format in questionnaires [Editorial]. *European Journal of Psychological Assessment, 34*(1), 1–5. https://doi.org/10.1027/1015-5759/a000469

Yu, M. V. B., Johnson, H. E., Deutsch, N. L., & Varga, S. M. (2018). "She calls me by my last name": Exploring adolescent perceptions of positive teacher-student relationships. *Journal of Adolescent Research, 33*(3), 332–362. https://doi.org/10.1177/0743558416684958

ORCID

Irene García-Moya
https://orcid.org/0000-0002-5377-315X

Irene García-Moya
Departamento de Psicología Evolutiva y de la Educación
C/Camilo José Cela s/n
41018 Seville
Spain
irenegm@us.es