Validation of the student version of the Perceptions of Inclusion Questionnaire on a sample of French students

Françoise Guillemot, Marco G.P. Hessels

Validation of the student version of the Perceptions of Inclusion Questionnaire on a sample of French students

Françoise Guillemot[a]*, Marco G.P. Hessels[b,c+]

[a]*Centre for Educational Research Nantes, Faculty of Psychology, Nantes, France;*

[b]*Faculty of Psychology and Educational Sciences, University of Geneva, Switzerland*

[c]*Optentia Research Focus Area, North-West University, Vanderbijlpark, South Africa*

*francoise.guillemot@univ-nantes.fr; https://orcid.org/0000-0001-8528-4142

[+] marco.hessels@unige.ch; https://orcid.org/0000-0002-4018-1374

Validation of the student version of the Perceptions of Inclusion Questionnaire on a sample of French students

The Perceptions of Inclusion Questionnaire (PIQ) measures emotional well-being at school, social inclusion, and academic self-concept of students aged 9 to 15 years, with and without Special Educational Needs (SEN). This study aimed to validate the PIQ with French students in 6th grade. The sample included 288 students with an average age of 11 years and 4 months. Confirmatory factor analyses supported construct validity of the 3-factors. Correlations with tests already validated in French were between .41 and .60, which demonstrate the convergent validity of the different domains of the PIQ. Internal consistency and test-retest reliability are quite satisfactory (Cronbach's $\alpha > .70$ and $r > .73$). The strict scalar invariance of the scale was verified according to gender. Furthermore, differential analyses according to gender and the presence of SEN revealed that girls show a slightly higher emotional well-being than boys and that students with SEN have a significantly lower academic self-concept than students without SEN. No other significant differences regarding gender and SEN were found. It was concluded that the French student-version of the PIQ has good psychometric properties and can be used as a simple tool for measuring inclusion of pupils with and without SEN.

Keywords: Inclusion; student perceptions; school well-being; social inclusion; academic self-concept; perceptions of inclusion questionnaire; special educational needs

Introduction

In line with the UN Convention on the Rights of Persons with Disabilities (United Nations, 2006), and as is the case for many countries worldwide, increasing numbers students with disabilities are enrolled in mainstream schools in France (Le Laidier, Michaudon, and Prouchandy 2016). In 2016-2017, 3.4% of students were officially diagnosed with a disability, of whom 1.4% were educated in non-inclusive environments (European Agency for Development in Special Needs Education 2020). Additionally, more and more students are recognised by their schools as having Special

Educational Needs (SEN) and benefit from special measures (adapted examinations, additional time, use of a computer) according to their difficulties. Currently, no statistics exist for the number of students with SEN in France, but it is estimated at less than 10% of all pupils.

International studies show significant benefits of inclusive education for students, but difficulties remain (Bear, Minke, and Manning 2002). Merely placing students with varying educational needs together in one classroom does not suffice for obtaining a truly inclusive school environment (Zurbriggen et al. 2017a). Various quality indicators have been proposed, covering school systems, teacher training and student results (Loreman, Forlin, and Sharma 2014; Kyriazopoulou and Weber 2009). In terms of outcomes, quality can be measured through school performance but also by assessing the students' subjective well-being, social integration and self-esteem at school. Thus, appropriate tools are needed that allow measuring these factors in students with and without SEN. The Perceptions of Inclusion Questionnaire (PIQ) is a recently developed tool that aims exactly the latter (Venetz et al. 2015). Our aim is a first validation of the PIQ student version in French.

Well-being at school, social inclusion and academic self-concept

Well-being at school

Perceived well-being at school can be defined as feeling good at school, sometimes also referred to as emotional inclusion, or subjective quality of life at school. Many questionnaires measure quality of life or subjective well-being in a global way, including some questions related to school, such as the Multidimensional Life Satisfaction Scale for Children (MLSS; Huebner 1994) and the KIDSCREEN-52 Quality-of-Life Measure for Children and Adolescents (Ravens-Sieberer et al. 2005).

Other questionnaires specifically aim school well-being in a single or multidimensional way, e.g., the Student Perceptions of Inclusion in Rural Canada (SPIRC; Loreman et al. 2008) or the French Well-being in primary and middle school (BE-SCOL; Guimard et al. 2015). A growing body of research shows that in high school, subjective well-being is related to academic performance and success in adult life (Gibbons and Silva 2011). However, no consensus exists regarding the outcomes of children with SEN. Some German and French studies show that mainstreamed children with SEN are equally happy in school as their peers without disabilities (Gebhardt et al. 2012; Godeau et al. 2015). Others, such as Stiefel et al. (2018), report a slightly lower level of well-being.

Social Inclusion

Students' social inclusion can be measured either by asking children directly about their perceived social inclusion or social self-concept, or indirectly, by asking other children, parents, or teachers about peer relationships. A scale frequently used is the French version of the Self-Perception Profile for Children (SPPC; Harter 1982). Although students with SEN report equally positive perceptions of their social relationships as their peers without disabilities (Avramidis, Avgeri, and Strogilos 2018; Koster et al. 2010; Zeleke 2004), their parents and teachers are more negative (Schwab, Zurbriggen, and Venetz 2020). Observations of interactions and popularity show that students with SEN have fewer friends and are less popular among their peers (Avramidis 2013; Koster et al. 2010). They are also more often victim of bullying (Rose, Monda-Amaya, and Espelage 2011). Furthermore, studies conducted with older students show a lower appreciation of social inclusion by students with SEN (Gebhardt et al. 2012; Schwab et al. 2016).

Academic self-concept

The self-concept refers to thoughts and feelings about oneself. Several scales have been developed to measure it in a general unidimensional way, like the Rosenberg scale (Rosenberg 1965), or multidimensionally with the SPCC (Harter 1982) or the Piers-Harris Children's Self-Concept Scale (Piers and Herzberg 2007). The academic self-concept is one dimension of the self-concept and corresponds to the student's image of the accomplishments in learning. It is positively related to academic performance. Nevertheless, the sense of causality is one of the most discussed topics in research on academic self-concept (Green et al. 2006) and a model of reciprocal effects seems most appropriate (Marsh and O'Mara 2008; Valentine, DuBois, and Cooper 2004; Valentine and DuBois 2005). A meta-analysis of 62 studies concluded that the academic self-concept of students with learning disabilities is lower than that of non-disabled students (Bear, Minke, and Manning 2002), but the meta-analysis by Elbaum (2002) explicates that it may depend on the school context.

The Perceptions of Inclusion Questionnaire

The PIQ (Venetz et al. 2015) is a short questionnaire with 12 items, consisting of simple sentences that are easily understandable for students with or without learning difficulties. Three domains are addressed (four items each): emotional well-being in school, social inclusion in class and the academic self-concept. The PIQ consists of a version for students, one for parents and one for teachers, all using the same wording, but adapted to the perspective of the person responding.

Its psychometric properties were demonstrated in a study with 823 German-speaking Swiss students (mean age 12 years), including 190 students with SEN (Zurbriggen et al. 2017a; 2017b). The PIQ proved to be a reliable tool, particularly at the lower levels

of the scales indicating limited inclusion. There is no bias due to exogenous variables - having SEN or not- except for one item (’I am a fast learner’) that shows a minor deviation. The PIQ has been extended for use with students in grades 3 to 9 (Venetz, Zurbriggen, and Schwab 2019). Subsequently, the teacher version (PIQ-T) was tested and the consistency between teachers and students was evaluated (Schwab and Alnahdi 2020; Venetz, Zurbriggen, and Schwab 2019). More recently, the parent version (PIQ-P) showed good psychometric qualities (Schwab, Zurbriggen, and Venetz 2020). The questionnaire has been validated for students with disabilities, but negatively worded items appeared to be more problematic (Knickenberg et al. 2019). Furthermore, structural and scalar invariance were shown for students with disabilities depending on the schooling context (Knickenberg et al. 2019). Although the PIQ is currently available in 24 languages, and in different countries studies are under way or have been published (e;g., Alnahdi and Schwab 2020; Trygger 2019), the validation studies mentioned concern only the German version of the PIQ.

The present study

Since the PIQ's psychometric properties cannot be considered universally valid, it is important to scrutinise the instrument's characteristics in each context in which it is applied. Thus, our aim is to investigate some of the psychometric properties of the French version of the PIQ, such as its factor structure and convergent validity. If the PIQ functions in the same way as the German language version, we should find the same factor structure, as well as positive correlations with similar scales in French. The effects of gender and being identified as having SEN are also tested in order to confirm the results of the German studies.

Materials and methods

Participants

The participants were sixth grade students from 10 different classes in the same school in Loire Atlantique (France). The parents of all students were invited by e-mail to agree to their child's participation in the study: 97% agreed. The final sample included 134 boys (46.5%) and 154 girls (53.5%), aged between 9;10 and 12;10 years (M=11;4 years, SD=4 months). Nineteen students (6.6%) were identified as having SEN by their school administration based on internal evaluations. Three of them (1%) had a formal recognition of their disability as specified by national law. Data were collected in November 2020.

Measures

The PIQ student version

The French version of the PIQ proposed by the authors (https://piqinfo.ch/sprachversionen/) comprises statements such as 'I like going to school'. Students are asked to indicate their agreement using a 4-point Likert scale ranging from 'not at all true' (1) to 'certainly true' (4). One statement in each dimension, represents a negative sentiment, for example, 'I feel alone in my class'. The score is calculated by averaging the scores of the four items of each dimension (after inverting the negatively worded items). Thus, a higher score represents greater emotional well-being in school, greater social inclusion and higher academic self-esteem.

School dimension of the MLSS

The short version of the MLSS comprises 30 items (Huebner, Zullig, and Saha 2012). The French version has been validated with children aged 5 to 11 years, with and

without intellectual disabilities (Coudronnière et al. 2017). To investigate convergent validity, emotional well-being of the PIQ was compared to the MLSS school domain, which comprises six items. The MLSS uses a 6-point Likert scale from ‘completely disagree’ (1) to ‘completely agree’ (6). A higher score indicates greater well-being in school.

Academic and social self-perception

The SPPC is a 36-item questionnaire that measures self-perceptions along 6 dimensions (Harter 1982). A validated version is available in French (Pierrehumbert, Plancherel, and Jankech-Caretta 1987). The presentation of the questionnaire is peculiar, in the sense that two groups of children are described, for instance: ‘some children work slowly’ and ‘other children do their work quickly’. The student has to indicate which description fits most, specifying whether these descriptions are ‘really true or ‘sort of true’. Such a presentation is claimed to reduce the positive/negative valence of the questions. The six items corresponding to competencies in the school domain were used to evaluate convergent validity with academic self-concept of the PIQ. The six items in the social domain were used for the social inclusion scale.

Peer Acceptance

Students were asked to name their best friends in class (maximum 5). This procedure allows to assess peer acceptance by means of the number of nominations each child received (Frederickson et al. 2007; Pijl and Frostad 2010).

Perception of bullying

Concerning bullying, six questions were asked, that were inspired on the Health Behaviour in School-aged Children questionnaire (HBSC; Godeau et al. 2015) and the California Bully Victimization Scale Scale (CBVS; Felix et al. 2011). The items

concern physical, moral and social bullying. We added cyber bullying which was not previously included. Students were asked to rate the frequency of different forms of harassment on a 4-point scale with anchors ‘never’ (1) and ‘weekly’ (4). A high score indicates a high level of bullying.

Procedure

The questionnaires were presented on computer during the science class in school. The researchers were present for the first four sessions with the teachers in order to train them. Thereafter, the teachers conducted the administration of the questionnaires on their own (two teachers per group of 12 students). The students were first asked to give their consent to participation in the research, after which they were asked to answer a first practice question. It was then assured that every single student had understood. Students who needed help reading or understanding a question could ask for assistance from the researchers present. Next, the PIQ questions appeared on the screen one by one and the students had to enter and validate their answer. The other questionnaires had the same format of presentation. Lastly, the students answered socio-demographic questions, including whether or not the student was receiving assistance because of SEN. Official data regarding SEN were also collected from the administration. A subsample of 52 students (two entire classes) participated in a second assessment that took place one week later, under the same conditions, to evaluate test-retest reliability. Students did not require help when answering to the PIQ and the MLSS scale was also completed without difficulty. The students did have some difficulties in grasping Harter's SPPC and how to position themselves, but were able to respond with the assistance provided.

Statistical analysis

To estimate the sample size needed for the confirmatory analyses, we follow the recommendations by Wolf et al (2013), taking the number of items per factor, as well as the item-loadings into account. According to the first validation study, there are 4 items per factor and the loadings are all higher than 0.8, (Zurbriggen et al. 2017a). In this situation, a sample size above 150 is required to guarantee a power greater than .80.

Multi-normality of responses to the 12 PIQ items was tested using Mardia's statistic (1970): data must be considered as non-normal when it exceeds a value of 5 (Bentler 2006). Next, a confirmatory analysis for categorical data was conducted (CCFA). The robust Weighted least square means and variances estimator (WSLMV) was preferred to a classical robust maximum likelihood estimator, given the non-normality of the categorical data (Li 2016) and the number of modalities of the likert scale which is less than 5 (Rhemtulla, Brosseau-Liard, and Savalei 2012) . Analyses were conducted with R software (R Core Team 2021), especially its Lavaan and SemTools package (Rosseel 2012; Jorgensen et al. 2021). Different indices were computed to evaluate the quality of the model. Since sample size was large, the χ^2 distance was expected to be generally significant and, thus, would not have allowed to evaluate model adequacy. TLI (Tucker Lewis Index) and CFI (Comparative Fit Index) values above .9 would indicate a satisfactory fit, and above .95 a good fit (Hu and Bentler 1999). The RMSEA (Root Mean Square Error of Approximation) indicator should be as close to 0 as possible, a value less than 0.05 indicates a good fit, and it is recommended that the confidence interval has an upper bound less than 0.08 (MacCallum, Browne, and Sugawara 1996). To improve the model, a study of index changes was conducted. This procedure indicates which paths need to be added to the initial model to reduce the χ^2 distance and thereby improve model fit. In addition, convergent validity was studied by calculating

the correlations between the 3 subscales of the PIQ and scales relating to emotional well-being at school (MLSS), perception of the social self-concept (social SPSS) and the academic self-concept (academic SPCC).

For reliability, Cronbach's α and McDonald's ω coefficients were calculated, with a value greater than .7 indicating satisfactory consistency (Nunnally and Bernstein 1994). Finally, test-retest reliability was assessed by calculating correlations between the first and second assessments and Student's tests for paired groups were calculated to exclude significant differences between assessments. A Bland Altman graph (Bland and Altman 1995) was plotted to check for heteroskedasticity or outliers (Aldridge, Dovey, and Wade 2017).

Invariance of measurement regarding gender was studied using a multi-group CCFA model using, as before, WLSMV estimating. We tested configural, metric, scalar and finally strict scalar invariance. To confirm measurement invariance, the chi-square test between two successive models should be not significant, and the difference of the CFI should be less than .01 (Cheung & Rensvold, 2002).

For differential analyses, simple Student's tests were performed to compare the different groups. Since the numbers considered are low for students with SEN, they did not allow for a study of measurement invariance or multiple regressions. The degrees of freedom were corrected to take heterogeneity of variances into account, in which case the results of the Welch's tests were given.

Results

Descriptive statistics

The means and standard deviations of the PIQ subscales, and the means and standard

deviations of the other measures are presented in **Erreur ! Source du renvoi introuvable.**. The three PIQ means are high, (M=3.17, 3.38, 2.94) with homogeneous standard deviations (SD=0.56, 0.51, 0.54). The means are above the midpoint of the scale, which is 2.5 for a scale ranging from 1 to 4. The MLSS school subscale has a very high mean of 5.17 for a scale ranging from 1 to 6, with a midpoint of 3.5. The SPPC means are also higher than the theoretical mean. The mean bullying score is 1.54 (SD=0.448), with a score of 1 indicating no bullying and 4 corresponding to strong bullying. The number of citations ranged from 0 (the student was never mentioned by his peers) to 9 (the student was mentioned by 9 different peers) with a mean of 3.14 (SD=1.96).

Factorial Validity

The first model tested corresponds to the model proposed by the authors of the PIQ. Items 1, 4(-), 7 and 10 belong to the first dimension, items 2, 5, 8(-) and 11 to the second and items 3, 6, 9 and 12(-) to the third. Fit indicators are good, χ^2 (51) = 134.3, p < .001, CFI = .944, TLI = .928, RMSEA = .075 (90% CI = [.060, .091] p = .004). The 12 loadings as well as the 3 estimated covariances are significant (p<.001), the values are shown on the arrows in Figure 1.

Two other models were tested: the model 2 with an additional path between item 1 and 4, the model 3 with another additional path between the dimension social inclusion and item 7. These two models show very good properties (see supplementary material).

Convergent validity

The MLSS school subscale and the two subscales from Harter's SPPC show good internal consistency in our sample. For the school dimension of the MLSS, Cronbach's α is .86 (McDonald's ω = .87), for the social dimension of the SPPC α = .64 (ω = .67)

and for the academic dimension of the SPPC $\alpha = .59$ ($\omega = .61$). The correlations between the PIQ subscales and the other scales are presented in **Erreur ! Source du renvoi introuvable.**.

The correlations are all moderate to high (above .40) and significant ($p < .001$) which shows a good convergent validity between the PIQ scales and the corresponding scales of the other tests. The correlation between the academic self-concept of the PIQ and the SPPC school domain is the highest ($r = .63$). Social inclusion of the PIQ and the SPPC social domain show a moderate correlation ($r = .41$). The correlation between emotional well-being of the PIQ and the corresponding MLSS subscale is also high ($r = .54$). The other correlations are lower, indicating good divergent validity. The convergent validity of the PIQ's social inclusion scale was further investigated using two additional measures, a bullying measure and a nomination-of-friends-procedure. A moderate negative correlation was found between social inclusion and the level of bullying ($r = -.42$, $p < .001$). The nomination-of-friends-procedure provides an external measure of social inclusion, as it is based on declarations by peers. Here, too, a moderate correlation was found between PIQ scores for the social inclusion and the number of times a student was nominated as a friend by peers ($r = .36$, $p < .001$).

Internal consistency and test-retest reliability

The indices of internal consistency, calculated for categorical data, are presented in **Erreur ! Source du renvoi introuvable.**. All are above .70, indicating good internal consistency (Nunnally and Bernstein 1994).

The correlations between the measures at T1 and T2 were calculated for the 52 students who completed the questionnaire twice (see **Erreur ! Source du renvoi introuvable.**). These were all greater than .70, indicating good test-retest reliability

(Chadha 2009). Student tests for paired groups did not show any significant difference between the measures at T1 and T2, except for social inclusion where a small, but significant decrease was observed ($t(51) = 2.020$, $p = .049$, Cohen's $d = .28$). Bland Altman's graphs were plotted, as well as confidence intervals for the approval limits. These confirmed that no differences existed between the two measures. Moreover, individual differences between T1 and T2 were always less than 0.75 points.

Gender measurement invariance

First of all, the TLI and CFI indices of the configural model (see **Erreur ! Source du renvoi introuvable.**) were acceptable (CFI = .921, TLI = .897). Then, we estimated the metric model and compared it to the configural model, which showed no significant differences between the two models ($\Delta\chi^2(9) = 15.36$, $p = .081$). Next, we estimated the scalar model, which did not differ significantly from the metric model ($\Delta\chi^2(19) = 25.70$, $p = .138$) and, finally, the strict scalar model, which also did not differ significantly from the scalar model ($\Delta\chi^2(12) = 14.33$, $p = .28$). Moreover, the differences in CFI were less than .01 (Cheung & Rensvold, 2002), which confirms strict scalar measurement invariance for the 3-factor model according to gender.

Differential analyses

Differential analyses according to gender showed only one significant difference between girls and boys (see Table 6): boys report a lower emotional well-being than girls, but the effect size is small ($t(286) = -2.228$, $p = .027$, $d = -.26$). Bullying measures and peer nominations did not show any differences between girls and boys.

Regarding the presence of SEN, no significant differences were observed for emotional well-being and social inclusion. For the academic self-concept, pupils with SEN showed significantly lower scores with a large effect size $t(286) = 3.26$, $p = .019$, d

= .78. The other measures of social inclusion, i.e. bullying and peer nominations, also showed significant differences between students with and without SEN, with high effect sizes. Pupils with SEN were bullied more often and were less often mentioned as a friend by their peers (one less friend on average, see Table 7).

Discussion

This study aimed to validate the French-language version of the PIQ questionnaire for 6th grade students. The first findings were that the means of the PIQ scales were relatively high, and well above the theoretical mid-point of the scale which is 2.5. This was especially true for emotional well-being and social inclusion, but less so for the academic self-concept. This confirms results from studies, for example, in Germany (Schwab and Alnahdi 2020). Students report to be happy at school, well integrated socially and show a high academic self-concept.

The confirmatory factor analysis confirmed the three-factor structure as proposed by Zurbriggen et al. (2017a). The addition of a relationship between items 1 and 4 of the first dimension improved the adjustment indices. This seems very logical since these two items, in French, appear as the negation of each other: ‘I like going to school’ and ‘I have no desire to go to school’. The second possible improvement is to add a path from the social domain to item 7 ‘I like it in school’. This improvement is identical to the one proposed in the Swedish language (Trygger 2019). Item 7 belongs to the emotional dimension but is also strongly related to social inclusion.

Convergent validity was established for emotional well-being of PIQ using the MLSS scale relative to school (Huebner, Zullig, and Saha 2012). Both scales show a very good correlation. For the social inclusion and academic self-concept, two subscales from the SPCC were used (Harter 1982). During assessment, responding to the SPCC

items proved to be more difficult for students. They asked several questions that concerned understanding of what was meant, as well as about the how the scale functioned. Despite these difficulties, the scale appeared reliable and correlations between the SPCC Academic Self-concept and the PIQ Academic Self-concept were high. The correlation was somewhat lower between the SPCC Social Self-concept and the PIQ Social Inclusion. The results of the present study concur with those obtained in the initial validation study of the PIQ. In the latter study, convergent validity was estimated by means of several different student and teacher measures (e.g., the Strengths and Difficulties Questionnaire (SDQ; Goodman 2001). Correlations were in the same order of magnitude as those found here (Zurbriggen et al. 2017b).

The reliability of the student version of the PIQ appeared to be satisfactory. Although Cronbach's α and McDonald's ω were all above .7, they were slightly lower than those reported by Zurbriggen et al. (2017). To our knowledge, test-retest reliability had not yet been studied. In our study, 52 participants completed the questionnaire twice, one week apart, under equal conditions. The correlations were high, indicating good test-retest reliability.

Strict scalar measurement invariance according to gender was demonstrated, which attests that the three concepts are measured equivalently for girls and boys and, consequently, allow comparisons between groups. Unfortunately, the small number of students with SEN in our study did not allow to evaluate invariance on this factor. A study by Knickenberg et al. (2019) demonstrated partial invariance for students with SEN, depending on the school context, especially regarding negatively formulated items.

Differential analyses showed that emotional well-being varied slightly by gender. This differs from DeVries, Voß, and Gebhardt (2018) who did not find any

gender differences but corresponds to the results found for French-speaking students in Switzerland (Beza 2018) and in earlier studies in the Netherlands (e.g., Verkuyten and Thijs 2002). Students with SEN show a weaker academic self-concept, which is consistent with the study of DeVries et al. (2018), as well as other studies using other questionnaires (Bear, Minke, and Manning 2002). Contrary to the results of DeVries et al. (2018), the emotional well-being of students with SEN in our sample was not significantly lower. The same result was found in France in a previous larger study (Godeau et al. 2015), and can be explained by a relatively inclusive school environment. In France almost all pupils attend the same schools until they are 16. With regard to social inclusion, pupils with SEN reported an equally high social inclusion as those without SEN, confirming earlier results (DeVries, Voß, and Gebhardt 2018; Venetz, Zurbriggen, and Schwab 2019). It is important to note that other measures of social inclusion did show significant differences that were not observed in the PIQ. Students with SEN reported being bullied more often, and are less frequently mentioned as a friend by their peers.

One of the limitations of our study is that only pupils in sixth grade participated in the study. Of course, it is necessary to be able to generalise the results to pupils between 8 and 16 years old. Therefore, samples with a larger age range and over a wider sample of schools need to be assessed. The small sample size did also not allow to study structural invariance between pupils with and without SEN; the number of students with SEN was very low in this study. Finally, parents and teachers were not yet included in the present research.

To conclude, the PIQ seems to be a very suitable tool to assess emotional well-being, social inclusion and the academic self-concept of French students. Compared to existing tools in French, the simplicity of the PIQ make it a favourable tool for studies involving

both students with and without SEN. Moreover, administration by means of computers might also be a way to obtain data that are more valid with students with SEN, for instance because presenting questions sequentially on a screen allows students to concentrate on one question at a time. For pupils with more severe disabilities, however, it might also be necessary to read the questions aloud and/or to provide pictorial materials and training in responding to Likert items, to favour understanding, as well as to avoid response bias (Hartley and MacLean 2006).

An interesting prospect is the possibility to use this tool in international comparisons, as it has been translated into more than 20 languages, and has been used in different countries. Nevertheless, it remains important to also include other measures, for example, a self-report of social inclusion needs to be completed by external measures, such as the friends-nomination-procedure, to detect difficulties that pupils would not necessarily express. A specific measure of bullying also appears useful. Together with the PIQ for parents and teachers (Venetz et al. 2015) a more comprehensive picture of inclusion and well-being would be obtained.

References

Aldridge, Victoria K., Terence M. Dovey, and Angie Wade. 2017. 'Assessing Test-Retest Reliability of Psychological Measures: Persistent Methodological Problems'. *European Psychologist* 22 (4): 207–18. https://doi.org/10.1027/1016-9040/a000298.

Alnahdi, Ghaleb H, and Susanne Schwab. 2020. 'Inclusive Education in Saudi Arabia and Germany: Students' Perception of School Well-Being, Social Inclusion, and Academic Self-Concept'. *European Journal of Special Needs Education*, September, 1–14. https://doi.org/10.1080/08856257.2020.1823163.

Avramidis, Elias. 2013. 'Self-Concept, Social Position and Social Participation of Pupils with SEN in Mainstream Primary Schools'. *Research Papers in Education* 28 (4): 421–42. https://doi.org/10.1080/02671522.2012.673006.

Avramidis, Elias, Georgia Avgeri, and Vasilis Strogilos. 2018. 'Social Participation and Friendship Quality of Students with Special Educational Needs in Regular Greek Primary Schools'. *European Journal of Special Needs Education* 33 (2): 221–34. https://doi.org/10.1080/08856257.2018.1424779.

Bear, George G., Kathleen M. Minke, and Maureen A. Manning. 2002. 'Self-Concept of Students with Learning Disabilities: A Meta-Analysis.' *School Psychology Review* 31 (3): 405–27.

Bentler, P. M. 2006. *EQS 6 Structural Equations Program Manual*. Encino, CA: Multivariate Software, Inc.

Beza, Maria. 2018. 'La participation sociale et le bien-être des élèves à Genève [Social participation and well-being of students in Geneva]'. University of Geneva. https://archive-ouverte.unige.ch/unige:110066.

Bland, J. M., and D. G. Altman. 1995. 'Comparing Methods of Measurement: Why Plotting Difference against Standard Method Is Misleading'. *Lancet (London, England)* 346 (8982): 1085–87. https://doi.org/10.1016/s0140-6736(95)91748-9.

Chadha, N. K. 2009. *Applied Psychometry*. London ; Thousand Oaks, CA: SAGE.

Coudronnière, Charlotte, Fabien Bacro, Philippe Guimard, and Jean-Baptiste Muller. 2017. 'Validation of a French Adaptation of the Multidimensional Student's Life Satisfaction Scale in Its Abbreviated Form, for 5- to 11-Year-Old Children with and without Intellectual Disability'. *Journal of Intellectual and Developmental Disability*, July, 1–14. https://doi.org/10.3109/13668250.2017.1285012.

DeVries, Jeffrey M., Stefan Voß, and Markus Gebhardt. 2018. 'Do Learners with Special Education Needs Really Feel Included? Evidence from the Perception of Inclusion Questionnaire and Strengths and Difficulties Questionnaire'. *Research in Developmental Disabilities* 83 (December): 28–36. https://doi.org/10.1016/j.ridd.2018.07.007.

Elbaum, Batya. 2002. 'The Self-Concept of Students with Learning Disabilities: A Meta-Analysis of Comparisons Across Different Placements'. *Learning Disabilities Research and Practice* 17 (4): 216–26. https://doi.org/10.1111/1540-5826.00047.

European Agency for Development in Special Needs Education. 2020. *European Agency Statistics on Inclusive Education: 2018 Dataset Cross-Country Report.* J. Ramberg, A. Lénárt and A. Watkins. Odense, Denmark.

Felix, Erika D., Jill D. Sharkey, Jennifer Greif Green, Michael J. Furlong, and Diane Tanigawa. 2011. 'Getting Precise and Pragmatic about the Assessment of Bullying: The Development of the California Bullying Victimization Scale'. *Aggressive Behavior* 37 (3): 234–47. https://doi.org/10.1002/ab.20389.

Frederickson, Norah, Elizabeth Simmonds, Lynda Evans, and Chris Soulsby. 2007. 'Assessing the Social and Affective Outcomes of Inclusion'. *British Journal of Special Education* 34 (2): 105–15. https://doi.org/10.1111/j.1467-8578.2007.00463.x.

Gebhardt, Markus, Susanne Schwab, Mathias Krammer, and Klicpera Gasteiger. 2012. 'Achievement and Integration of Students with and without Special Educational Needs (SEN) in the Fifth Grade'. *Journal of Special Education and Rehabilitation* 13 (3–4). https://doi.org/10.2478/v10215-011-0022-6.

Gibbons, Stephen, and Olmo Silva. 2011. 'School Quality, Child Wellbeing and Parents' Satisfaction'. *Economics of Education Review* 30 (2): 312–31. https://doi.org/10.1016/j.econedurev.2010.11.001.

Godeau, Emmanuelle, Mariane Sentenac, D Pacorina, and V Ehlinger. 2015. 'Élèves handicapés ou porteurs de maladies chroniques : perception de leur vie et de leur bien-être au collège. [Students with disabilities or chronic illnesses: perceptions of their life and well-being at college.]', Éducation et formations, , no. 88–89 (December): 145–61.

Goodman, Robert. 2001. ‘Psychometric Properties of the Strengths and Difficulties Questionnaire’. *Journal of the American Academy of Child & Adolescent Psychiatry* 40 (11): 1337–45. https://doi.org/10.1097/00004583-200111000-00015.

Green, Jasmine, Genevieve Nelson, Andrew J. Martin, and Herb Marsh. 2006. ‘The Causal Ordering of Self-Concept and Academic Motivation and Its Effect on Academic Achievement’. *International Education Journal* 7 (4): 534–46.

Guimard, Philippe, Fabien Bacro, Séverine Ferrière, Agnès Florin, Tiphaine Gaudonville, and Hué Ngo. 2015. ‘Le bien-être des élèves à l’école et au collège. Validation d’une échelle multidimensionnelle, analyses descriptives et différentielles. [Students’ well-being at school and college. Validation of a multidimensional scale, descriptive and differential analyses]’, 23.

Harter, Susan. 1982. ‘The Perceived Competence Scale for Children’. *Child Development* 53 (1): 87. https://doi.org/10.2307/1129640.

Hartley, S. L., and W. E. MacLean. 2006. ‘A Review of the Reliability and Validity of Likert-Type Scales for People with Intellectual Disability’. *Journal of Intellectual Disability Research* 50 (11): 813–27. https://doi.org/10.1111/j.1365-2788.2006.00844.x.

Hu, Li-tze, and Peter M. Bentler. 1999. ‘Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives’. *Structural Equation Modeling: A Multidisciplinary Journal* 6 (1): 1–55. https://doi.org/10.1080/10705519909540118.

Huebner, E. Scott. 1994. ‘Preliminary Development and Validation of a Multidimensional Life Satisfaction Scale for Children.’ *Psychological Assessment* 6 (2): 149–58. https://doi.org/10.1037/1040-3590.6.2.149.

Huebner, E. Scott, Keith J. Zullig, and Runa Saha. 2012. ‘Factor Structure and Reliability of an Abbreviated Version of the Multidimensional Students’ Life Satisfaction Scale’. *Child Indicators Research* 5 (4): 651–57. https://doi.org/10.1007/s12187-012-9140-z.

Jorgensen, Terrence D., Sunthud Pornprasertmanit, Alexander M. Schoemann, and Yves Rosseel. 2021. *SemTools: Useful Tools for Structural Equation Modeling*. https://CRAN.R-project.org/package=semTools.

Knickenberg, Margarita, Carmen L. A. Zurbriggen, Martin Venetz, Susanne Schwab, and Markus Gebhardt. 2019. ‘Assessing Dimensions of Inclusion from Students’ Perspective – Measurement Invariance across Students with Learning Disabilities in Different Educational Settings’. *European Journal of Special Needs Education*, July, 1–16. https://doi.org/10.1080/08856257.2019.1646958.

Koster, Marloes, Sip Jan Pijl, Han Nakken, and Els Van Houten. 2010. ‘Social Participation of Students with Special Needs in Regular Primary Education in the Netherlands’. *International Journal of Disability, Development and Education* 57 (1): 59–75. https://doi.org/10.1080/10349120903537905.

Kyriazopoulou, Mary, and Harald Weber. 2009. *Development of a Set of Indicators – for Inclusive Education in Europe*. Odensee, Denmark: European Agency for Development in Special Needs Education.

Le Laidier, Sylvie, Hélène Michaudon, and Patricia Prouchandy. 2016. ‘Depuis la loi de 2005, la scolarisation des enfants en situation de handicap a très fortement progressé. [Since the 2005 law, the schooling of children with disabilities has progressed considerably].’ *Note d’information MENESR-DEPP*, MENESR-DEPP, , no. 16.36 (December).

Li, Cheng-Hsien. 2016. 'Confirmatory Factor Analysis with Ordinal Data: Comparing Robust Maximum Likelihood and Diagonally Weighted Least Squares'. *Behavior Research Methods* 48 (3): 936–49. https://doi.org/10.3758/s13428-015-0619-7.

Loreman, Tim, Chris Forlin, and Umesh Sharma. 2014. 'Measuring Indicators of Inclusive Education: A Systematic Review of the Literature'. *International Perspectives on Inclusive Education* 3 (October): 165–87. https://doi.org/10.1108/S1479-363620140000003024.

Loreman, Tim, Judy Lupart, Donna McGhie-Richmond, and Jen Barber. 2008. 'The Development of a Canadian Instrument for Measuring Student Views of Their Inclusive School Environment in a Rural Context: The Student Perceptions of Inclusion in Rural Canada (Spirc) Scale.' *International Journal of Special Education* 23 (3).

MacCallum, Robert C., Michael W. Browne, and Hazuki M. Sugawara. 1996. 'Power Analysis and Determination of Sample Size for Covariance Structure Modeling.' *Psychological Methods* 1 (2): 130–49. https://doi.org/10.1037/1082-989X.1.2.130.

Mardia, K. V. 1970. 'Measures of Multivariate Skewness and Kurtosis with Applications'. *Biometrika* 57 (3): 519–30. https://doi.org/10.1093/biomet/57.3.519.

Marsh, Herbert W., and Alison O'Mara. 2008. 'Reciprocal Effects Between Academic Self-Concept, Self-Esteem, Achievement, and Attainment Over Seven Adolescent Years: Unidimensional and Multidimensional Perspectives of Self-Concept'. *Personality and Social Psychology Bulletin* 34 (4): 542–52. https://doi.org/10.1177/0146167207312313.

Nunnally, Jum C., and Ira H. Bernstein. 1994. *Psychometric Theory*. 3rd ed. McGraw-Hill Series in Psychology. New York: McGraw-Hill.

Pierrehumbert, B., B. Plancherel, and C. Jankech-Caretta. 1987. 'Image de Soi et Perception Des Compétences Propres Chez l'enfant: Présentation d'un Questionnaire Récent d'estime de Soi Pour Enfants. [Self-Image and Sense of Competence in Children: Presentation of a Recent Self-Esteem Questionnaire for Children.]'. *Revue de Psychologie Appliquée* 37 (4): 359–77.

Piers, E.V., and D.S. Herzberg. 2007. *Piers-Harris Children's Self-Concept Scale*. Second Edition. L.A.: WPS.

Pijl, Sip Jan, and Per Frostad. 2010. 'Peer Acceptance and Self-concept of Students with Disabilities in Regular Education'. *European Journal of Special Needs Education* 25 (1): 93–105. https://doi.org/10.1080/08856250903450947.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Ravens-Sieberer, Ulrike, Angela Gosch, Luis Rajmil, Michael Erhart, Jeanet Bruil, Wolfgang Duer, Pascal Auquier, et al. 2005. 'KIDSCREEN-52 Quality-of-Life Measure for Children and Adolescents'. *Expert Review of Pharmacoeconomics & Outcomes Research* 5 (3): 353–64. https://doi.org/10.1586/14737167.5.3.353.

Rhemtulla, Mijke, Patricia É Brosseau-Liard, and Victoria Savalei. 2012. 'When Can Categorical Variables Be Treated as Continuous? A Comparison of Robust Continuous and Categorical SEM Estimation Methods under Suboptimal Conditions'. *Psychological Methods* 17 (3): 354–73. https://doi.org/10.1037/a0029315.

Rose, Chad A., Lisa E. Monda-Amaya, and Dorothy L. Espelage. 2011. 'Bullying Perpetration and Victimization in Special Education: A Review of the Literature'. *Remedial and Special Education* 32 (2): 114–30. https://doi.org/10.1177/0741932510361247.

Rosenberg, M. 1965. *Society and the Adolescent Self-Image*. Princeton: NJ: Princeton University Press.

Rosseel, Yves. 2012. 'Lavaan: An R Package for Structural Equation Modeling'. *Journal of Statistical Software* 48 (2): 1–36.

Schwab, Susanne, and Ghaleb Hamad Alnahdi. 2020. 'Teachers' Judgments of Students' School-Wellbeing, Social Inclusion, and Academic Self-Concept: A Multi-Trait-Multimethod Analysis Using the Perception of Inclusion Questionnaire'. *Frontiers in Psychology* 11 (July). https://doi.org/10.3389/fpsyg.2020.01498.

Schwab, Susanne, Markus Gebhardt, Marco G.P. Hessels, and Lena Nusser. 2016. 'Predicting a High Rate of Self-Assessed and Parent-Assessed Peer Problems—Is It Typical for Students with Disabilities?' *Research in Developmental Disabilities* 49–50 (February): 196–204. https://doi.org/10.1016/j.ridd.2015.11.026.

Schwab, Susanne, Carmen L.A. Zurbriggen, and Martin Venetz. 2020. 'Agreement among Student, Parent and Teacher Ratings of School Inclusion: A Multitrait-Multimethod Analysis'. *Journal of School Psychology* 82 (October): 1–16. https://doi.org/10.1016/j.jsp.2020.07.003.

Stiefel, Leanna, Menbere Shiferaw, Amy Ellen Schwartz, and Michael Gottfried. 2018. 'Who Feels Included in School? Examining Feelings of Inclusion Among Students With Disabilities'. *Educational Researcher* 47 (2): 105–20. https://doi.org/10.3102/0013189X17738761.

Trygger, Maria. 2019. 'Do I Feel Included? A Study of Secondary Pupils' Perception of Inclusion in a Compulsory School in Sweden -with a Particular Focus on Pupils with Self-Reported Special Education Needs |'. https://www.researchgate.net/publication/334126293_Do_I_feel_included_A_study_of_Secondary_Pupils%27_Perception_of_Inclusion_in_a_Compulsory_School_in_Sweden_-with_a_particular_focus_on_pupils_with_self-reported_special_education_needs.

Valentine, Jeffrey C., and David L. DuBois. 2005. 'Effects of Self-Beliefs on Academic Achievement and Vice-Versa: Separating the Chicken from the Egg'. In *New Frontiers for Self Research*, edited by Herbert W. Marsh, Rhonda G. Craven, and D. M. McInerney, 53–78. International Advances in Self Research, v. 2. Greenwich, Conn: Information Age Pub.

Valentine, Jeffrey C., David L. DuBois, and Harris Cooper. 2004. 'The Relation Between Self-Beliefs and Academic Achievement: A Meta-Analytic Review'. *Educational Psychologist* 39 (2): 111–33. https://doi.org/10.1207/s15326985ep3902_3.

Venetz, Martin, Carmen L. A. Zurbriggen, Michael Eckhart, Susanne Schwab, and Marco G. P. Hessels. 2015. 'The Perceptions of Inclusion Questionnaire (PIQ). Version Française'. 2015. www.piqinfo.ch.

Venetz, Martin, Carmen L. A. Zurbriggen, and Susanne Schwab. 2019. 'What Do Teachers Think About Their Students' Inclusion? Consistency of Students' Self-Reports and Teacher Ratings'. *Frontiers in Psychology* 10 (July). https://doi.org/10.3389/fpsyg.2019.01637.

Verkuyten, Maykel, and Jochem Thijs. 2002. ‘School Satisfaction of Elementary School Children: The Role of Performance, Peer Relations, Ethnicity and Gender’. *Social Indicators Research* 59 (2): 203–28. https://doi.org/10.1023/A:1016279602893.

Wolf, Erika J., Kelly M. Harrington, Shaunna L. Clark, and Mark W. Miller. 2013. ‘Sample Size Requirements for Structural Equation Models: An Evaluation of Power, Bias, and Solution Propriety’. *Educational and Psychological Measurement* 73 (6): 913–34. https://doi.org/10.1177/0013164413495237.

Zeleke, Seleshi. 2004. ‘Self-concepts of Students with Learning Disabilities and Their Normally Achieving Peers: A Review’. *European Journal of Special Needs Education* 19 (2): 145–70. https://doi.org/10.1080/08856250410001678469.

Zurbriggen, Carmen L. A., Martin Venetz, Susanne Schwab, and Marco G. P. Hessels. 2017a. ‘A Psychometric Analysis of the Student Version of the Perceptions of Inclusion Questionnaire (PIQ)’. *European Journal of Psychological Assessment*, December, 1–9. https://doi.org/10.1027/1015-5759/a000443.

———. 2017b. ‘Electronic Supplementary Material 1 - Validity of the Perceptions of Inclusion Questionnaire (PIQ)’. http://econtent.hogrefe.com/doi/10.1027/1015-5759/a000443.

	PIQ_Emotional	PIQ_Social	PIQ_Academic	MLSS	SPPC_School	SPPC_Social	CBVS	NbNom
N	288	288	288	288	286	284	288	288
Missing	0	0	0	0	2	4	0	0
Mean	3.17	3.38	2.94	5.17	2.86	3.09	1.54	3.14
Standard deviation	0.56	0.51	0.54	1.01	0.51	0.53	0.45	1.96
Minimum	1	1.25	1	1	1.33	1.5	1.2	0
Maximum	4	4	4	6	4	4	4	9

Table 1. Descriptive statistics

	MLSS	SPPC _Social	SPPC _School	PIQ_Emotional	PIQ_Social	PIQ_Academic
MLSS						
SPPC _Social	.21					
SPPC _School	.30	.43				
PIQ_Emotional	**.54**	.19	.34			
PIQ_Social	.29	**.41**	.25	.32		
PIQ_Academic	.28	.27	**.63**	.34	.25	

All correlations $p \leq .001$. Values in bold concern convergent validity.

Table 2. Pearson correlations between the PIQ scales and the MLSS and SPPC scales

Reliability Statistics

	Cronbach's α	McDonald's ω
PIQ_Emotional	.74	.76
PIQ_Social	.73	.73
PIQ_Academic	.72	.73

Table 3. Indices of internal consistency of the PIQ subscales

	Pearson r	p	Paired Sample T-test (df=51)	p
PIQ_EmotionalT1 - PIQ_EmotionalT2	0.87	<.001	-0.882	0.382
PIQ_SocialT1- PIQ_SocialT2	0.73	<.001	2.020	0.049
PIQ_AcademicT1- PIQ_AcademicT2	0.82	<.001	-0.964	0.339

Table 4. Test-retest correlations and Student tests for paired groups (N=52)

	Df	Δ.Df	χ^2	Δ.χ^2	Pr(>χ^2)	cfi	Δ.cfi	rmsea	Δ.rmsea
config	102	-	184.29	-	-	.921	-	.095	-
metric	111	9	204.50	15.36	.081	.920	-.0005	.091	-.0036
scalar	130	19	221.60	25.70	.138	.914	-.0059	.087	-.0039
strict	142	12	234.68	14.33	.280	.915	.0010	.083	-.0043

Table 5. Gender invariance measure

		PIQ_Emotional	PIQ_Social	PIQ_Academic	CBVS	NbNom
Mean (SD)	Male (n=134)	3.09 (0.56)	3.37 (0.53)	2.93 (0.53)	1.60 (0.49)	3.02 (1.99)
Mean (SD)	Female (n=154)	3.24 (0.55)	3.38 (0.50)	2.95 (0.55)	1.49(0.40)	3.25 (1.94)
Student's t or Welch's t		t(286)=-2.228	t(286)=-0.138	t(286)=-0.355	t(254.9)=1.963	t(286)=0.969
p		0.027	0.890	0.723	0.051	0.334
Cohen's d		-0.26	-0.0163	-0.04	0.23	-0.11

Table 6. Effect of gender on the 3 subscales of PIQ, on bullying (CBVS) and on the number of nominations.

		PIQ_Emotional	PIQ_Social	PIQ_Academic	CBVS	NbNom
Mean (SD)	Not SEN (n=269)	3.18 (0.57)	3.40 (0.49)	2.96 (0.52)	1.52 (0.44)	3.22 (1.95)
Mean (SD)	SEN (n=19)	3.03 (0.46)	3.12 (0.68)	2.55 (0.68)	1.82 (0.44)	2.11 (1.85)
Student's t or Welch's t		t(286)=1.152	t(19.4)=1.749	t(286)=3.264	t(286)=-2.837	t(286)=2.406
p		0.250	0.096	0.001	0.005	0.017
Cohen's d		0.27	0.47	0.78	-0.67	0.57

Table 7. Impact of SEN on PIQ scales, bullying (CBVS), and number of nominations.

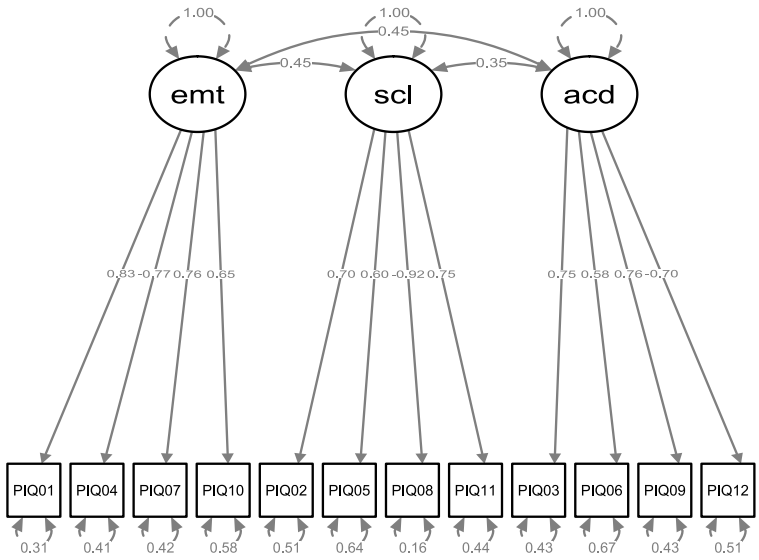

Figure 1. Model corresponding to the original factor structure