Evaluating the social outcomes of inclusive education
Koster, Marloes

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2008

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 09-11-2019
CHAPTER 5

EXAMINATION OF THE PSYCHOMETRIC QUALITIES OF THE SOCIAL PARTICIPATION QUESTIONNAIRE

A slightly adapted and shortened version of this article has been re-submitted for publication and was co-authored by Marieke E. Timmerman, Han Nakken, Sip Jan Pijl and Els J. van Houten
Abstract

Enhancing social participation of pupils with special needs is an important goal of inclusive education. This study addresses the psychometric qualities of a new teacher questionnaire to assess the social participation of pupils with special needs in regular primary education. This Social Participation Questionnaire initially consisted of 34 statements related to four key themes of social participation: ‘friendships/relationships’, ‘contacts/interactions’, ‘pupil’s social self-perception’ and ‘acceptance by classmates’, yielding four subscales.

A non-parametric item response analysis, a Mokken Scale Analysis, was used to examine the quality of the Questionnaire. Based on the analysis results, ten statements were removed. The resulting four subscales appeared intermediate to strong. As Mokken’s double monotonicity model turned out to fit well for each subscale, the subscale scores are on an ordinal scale and the separate statements are invariantly ordered. The subscale scores are comparable across pupils with and without special needs, because differential item functioning appeared to be absent. Subsequent analyses supported the division of statements into the four subscales.

The Questionnaire as a whole and its subscales were found to be reliable. Finally, evidence has been obtained for the Questionnaire’s discriminant validity, as the expected score differences between pupils with and without special needs were clearly present.
5.1 Introduction

Peer relationships are important in the live of children. It's widely acknowledged that childrens’ peer relationships and friendships contribute to their social and emotional development (Bagwell, 2004; Male, 2007) and also peer acceptance and social status are seen as vital to the well-being of children. Most primary school pupils find no difficulty in making friends and building positive relations with their peers (Powless & Elliott, 1993). However, this does not apply to pupils with special needs, as research has shown that these pupils experience considerably more difficulty building friendships in inclusive classrooms (Frostad & Pijl, 2007). This is striking, as within inclusive education social participation is considered to be vital. Parents often report increased social opportunities as their main motive for sending their child with special needs to a regular school (Sloper & Tyler, 1992). They expect their child to build positive relationships with typical peers. However, research has repeatedly shown that inclusion of pupils with special needs does not automatically lead to an increase of friendships between these pupils and their typical counterparts (De Monchy, Pijl & Zandberg, 2004). A study by Frostad and Pijl (2007) of Norwegian inclusive classrooms suggests that nearly a quarter of pupils with special needs have serious difficulties forming relationships in their peer group, while for their typical peers this is only 8 percent. This is worrisome, because the consequences of negative social experiences in school can be far-reaching. For instance, a low social status in childhood and peer relationship difficulties might lead to externalising (e.g. aggression) and internalising problems (e.g. anxiety) (Durrant, Cunningham & Voelker, 1990), and to maladjustment later in life (Bagwell, Newcomb & Bukowski, 1998; Parker & Asher, 1987). Because of the rather harmful long-term effects of negative social experiences at school, it is important to monitor the social participation of pupils. This entails a vital task for teachers.

Koster, Nakken, Pijl, Van Houten and Lutje Spelberg (in press) developed a questionnaire for teachers to assess the social participation of pupils with special needs. First, an analysis of literature was conducted in order to define the concept of social participation. This analysis revealed four key themes in social participation: friendships/relationships, contacts/interactions, pupil’s social self-perception and acceptance by classmates (Koster, Nakken, Pijl & Van Houten, in
press). The key themes laid the foundation for the so-called Social Participation Questionnaire, which aims at helping teachers make more accurate assessments of the social participation of their pupils, and aids in noticing problems in time.

In the present study, the quality of the Social Participation Questionnaire is central. Using empirical data from both pupils with and without special needs, the quality of the separate statements for each subscale, the separability of the four subscales and their reliability will be assessed. The Questionnaire’s discriminant validity will be examined by comparing the scores of pupils with special needs with those of pupils without special needs.

5.2 Method

5.2.1 Participants

For practical reasons, data collection took place in two periods. In each period, 300 regular primary schools were invited to participate in the study. According to files from the Ministry of Education, the invited schools have at least one pupil in Grades 1 to 3 who receive a so-called pupil-bound budget. Such a budget is allocated to pupils who, according to independent committees, meet the Dutch national criteria for a pupil-bound budget. Herewith, several categories of disabilities are distinguished, each with their own criteria. Those involve, among other things, categories of the Diagnostic and Statistical Manual of Mental Disorders (DSM IV) and intellectual ability, which are assessed by qualified psychiatrists or psychologists who operate independently from the committees. The invited schools were randomly drawn from the population of 2074 Dutch regular primary schools with a pupil with a pupil-bound budget in Grade 1, 2 and/or 3 within 2.5 hours of travelling time from the city of Groningen. This area, covering nearly two-third of the Netherlands, was chosen to make the data collection feasible.

First sub-sample. Of the 300 invited schools, 53 were involved in the study. One school was too busy to participate and 22 schools were unable to participate because they had no pupils with special needs in Grades 1, 2 or 3. The remaining 224 schools did not respond to our invitation. To examine possible bias in results from the cooperating schools, random selections of about 20 percent of the
schools who had not responded at all to our invitation to cooperate in the study (45 schools) and 35 percent of the schools who did cooperate (20 schools) were invited to complete a short questionnaire by phone.

This questionnaire is largely based on a questionnaire developed by Frostad and Pijl (2007). In total, 41 non-cooperating schools (91%) and 19 (95%) cooperating schools answered the questions. For the non-cooperating schools, the first question aimed at the reason for not participating in the study. It turned out that 12 schools were not able to participate, because they had no pupils with special needs in Grade 1, 2 or 3. These 12 schools were not asked to complete the remainder of the questionnaire. The remaining 29 non-cooperating schools and 19 cooperating schools completed the questionnaire. The questionnaire consisted of eight questions addressing attitudes to inclusion, teacher characteristics, satisfaction about the budget for materials and schooling, satisfaction about collaboration with parents and the degree to which teachers felt supported by the director, colleagues and peripatetic teachers at educating pupils with special needs. The questions had to be answered on a 5-point scale, ranging from 0 (yes, very much) to 4 (no, not at all). Comparison of the mean scores of cooperating and non-cooperating schools showed only slight differences, which appeared to be nonsignificant (see Table 1).
Table 1. Outcomes non-response survey first sub-sample

<table>
<thead>
<tr>
<th>Statements non-response questionnaire</th>
<th>Mean cooperating schools (n=19)</th>
<th>Mean non-cooperating schools (n=29)</th>
<th>95% CI dif</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attitude towards inclusion</td>
<td>1.05 (SD=0.68)</td>
<td>1.03 (SD=0.62)</td>
<td>-0.41;0.37</td>
</tr>
<tr>
<td>Support by director</td>
<td>0.89 (SD=0.74)</td>
<td>0.55 (SD=0.57)</td>
<td>-0.72;0.06</td>
</tr>
<tr>
<td>Support by colleagues</td>
<td>0.53 (SD=0.70)</td>
<td>0.66 (SD=0.61)</td>
<td>-0.26;0.51</td>
</tr>
<tr>
<td>Support by peripatetic teachers</td>
<td>0.42 (SD=0.51)</td>
<td>0.48 (SD=0.63)</td>
<td>-0.29;0.41</td>
</tr>
<tr>
<td>Satisfaction about collaboration with parents</td>
<td>0.53 (SD=1.12)</td>
<td>0.55 (SD=1.06)</td>
<td>-0.62;0.67</td>
</tr>
<tr>
<td>Satisfaction about budget for special materials</td>
<td>0.68 (SD=0.58)</td>
<td>0.72 (SD=0.59)</td>
<td>-0.31;0.39</td>
</tr>
<tr>
<td>Satisfaction about opportunities for schooling</td>
<td>0.89 (SD=0.32)</td>
<td>0.83 (SD=0.38)</td>
<td>-0.28;0.15</td>
</tr>
<tr>
<td>Degree to which extra schooling has taken place</td>
<td>1.79 (SD=0.86)</td>
<td>1.86 (SD=0.83)</td>
<td>-0.44;0.57</td>
</tr>
</tbody>
</table>

In addition, the outcomes showed that the mean total number of pupils on the schools, the mean number of pupils with special needs on the schools and the mean degree of teaching experience of teachers did not significantly differ for both groups of schools (total number of pupils: 95% CI of mean difference [95% CI dif]=-1.34;1.18; number of pupils with special needs: 95% CI dif=-2.1;3.78; degree of teaching experience: 95% CI dif=-1.56;1.18). The comparison suggests that the cooperating schools do not over-represent schools with the most positive view on and most positive experiences with inclusive education in the Netherlands. These outcomes suggest that there is no reason to assume any systematic bias in the sample involved in the study.

A final remark with regard to the sample concerns the noteworthy fact that almost 30 percent of the invited schools were unable to participate because they had no pupils with special needs in Grades 1 to 3. The names of the schools were distracted from files of the Ministry of Education. The data in the files applied to the school year 2005-2006, since there were no such files available for the school year 2006-2007 in which the study took place. Due to the fact that the
As described above, 53 schools participated in the study. Since several schools had a pupil with special needs in more than one class and since 15 classes contained two or more pupils with special needs, 75 classes with in total 96 pupils with special needs participated in the study. Next to the 96 pupils with special needs, the classes comprised 1652 pupils without special needs.

Second sub-sample. Of the 300 invited schools, 66 took part. Contrary to the first sub-sample, the schools of the second sub-sample were asked to respond to our invitation to participate in the study not only when they were willing to participate, but also when they did not want to participate. If schools did not want to take part in the study, they were asked to write down the reason for not participating. In total, more than 70 percent of the schools (n=211) responded to our invitation to participate. Of these 211 schools, 66 schools were willing to take part in the study, while 145 schools were not able or not willing to participate. Several reasons for not participating were mentioned. Almost 30 percent of the schools had no pupils with special needs in Grades 1 to 3, about 20 percent of the schools were too busy, and another 14 percent of the schools considered the study as being a too large burden for the pupils with special needs and their classmates. About 7 percent of the schools already participated in another study and almost 5 percent of the schools were not interested in the study. Several other schools (about 3 percent) were occupied with organisational changes, or parents did not gave permission to participate in the study (about 3 percent). The remainder of the schools (about 18 percent) did not mention a reason for not participating. As in the former sample the participating schools did not differ from the non-participating ones, there was no reason to assume any systematic bias in the sample involved in the study. Therefore, it was decided to refrain from repeating the non-response survey in the second sub-sample.

As named above, 66 schools took part in the study. In total, 27 schools had a pupil with special needs in more than one class and 34 classes contained two or more pupils with special needs. Consequently, 105 classes with in total 141 pupils with special needs participated in the study. The total number of pupils (including the 141 pupils with special needs and their classmates) in the 105 classes was 2426.
Total sample (sub-sample 1 and sub-sample 2). In total, 119 schools with 180 participating classes were involved. Each teacher was asked to fill in the Social Participation Questionnaire for the pupil(s) with special needs and for two pupils without special needs. As the Questionnaire was expected to be filled in for a diversity of children, this resulted in a broad range of scores. In seven classes, the teacher filled in the Questionnaire for two pupils with special needs and for only one pupil without special needs. The final sample thus consisted of 590 pupils, 237 of them with special needs and 353 without special needs. An overview of categories of disabilities and its distribution is presented in the second column of Table 2.

Table 2. Distribution of pupils with special needs into categories of disabilities and gender

<table>
<thead>
<tr>
<th>Category of disabilities</th>
<th>Number of pupils in absolute numbers and percentages</th>
<th>Boy</th>
<th>Girl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavioural disorder</td>
<td>29 (12.2%)</td>
<td>25 (86.2%)</td>
<td>4 (13.8%)</td>
</tr>
<tr>
<td>Autistic spectrum disorder</td>
<td>97 (40.9%)</td>
<td>83 (85.6%)</td>
<td>14 (14.4%)</td>
</tr>
<tr>
<td>Motor disability</td>
<td>35 (14.8%)</td>
<td>27 (77.1%)</td>
<td>8 (22.9%)</td>
</tr>
<tr>
<td>Intellectual disability</td>
<td>26 (11.0%)</td>
<td>10 (38.5%)</td>
<td>16 (61.5%)</td>
</tr>
<tr>
<td>Speech/language disabilities</td>
<td>47 (19.8%)</td>
<td>32 (68.1%)</td>
<td>15 (31.9%)</td>
</tr>
<tr>
<td>Learning disabilities</td>
<td>2 (0.8%)</td>
<td>2 (100%)</td>
<td>0</td>
</tr>
<tr>
<td>Chronic illness</td>
<td>1 (0.4%)</td>
<td>1 (100%)</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>237 (100%)</td>
<td>180 (75.9%)</td>
<td>57 (24.1%)</td>
</tr>
</tbody>
</table>

5.2.2 Instrument

The Social Participation Questionnaire was constructed during a step by step process. An exploration of literature (Koster, Nakken, Pijl & Van Houten, in press) revealed four key themes of social participation: friendships/relationships, contacts/interactions, pupil’s social self-perception and classmates’ acceptance. In addition, on the basis of literature, tens of statements concerning social participation were described. The content of the statements is mainly based on literature and partly on the clinical and research experience of the researchers who work for many years in the tradition of research looking at the social dimension of inclusion in education. Since this process might have been coloured by the researchers’ views, a panel consisting of 14 people with much practical experience in inclusive education was asked to critically examine the statements. The panel members were asked whether the formulated statements were in line
with their views of social participation, and to mention omissions. This detailed verification of the statements was intended to increase the chance of including all essential topics of social participation. Based on the panel’s comments, a list of 74 statements was compiled, yielding a picture of social participation (Koster, Nakken, Pijl, Van Houten & Lutje Spelberg, in press). A second check of the relevance of each statement was made using the ratings of a large group of respondents (n=190) who all had experience with inclusive education. The respondents were asked to rate the importance of each statement. Those statements rated as important or very important by at least 75 percent of the respondents were selected for the Questionnaire. In a pilot study the 30 resulting statements were psychometrically examined on the basis of scores obtained on 60 pupils (25 with special needs and 35 without). The results of a confirmatory factor analysis (Oblique Multiple Group Method, Holzinger, 1944) largely supported the Questionnaire’s division into four subscales. The analysis revealed that 22 out of 30 statements correlated strongest with the subscale they were assigned to. The remaining eight statements had higher correlations with one or more subscales they were not part of. The reliabilities of the ‘Friendships/relationships’ and ‘Contacts/interactions’ subscales were high (respectively 0.94 and 0.88), whereas the reliabilities of the ‘Social self-perception of pupil’ and ‘Acceptance by classmates’ subscales were fairly low (respectively 0.69 and 0.64). In order to further improve the reliability of the latter subscales, respectively three and two statements were added. One statement, belonging to the ‘Contacts/interactions’ subscale was removed from the list, as this statement was incongruous and correlated negatively with the total score on the Questionnaire. Thus, the revised Questionnaire under study here consisted of 34 statements representing the four key themes of social participation in 6 to 14 statements, which have to be scored on a 5-point scale (ranging from ‘this does not apply at all’ to ‘this strongly applies’). The majority of the statements (n=24) are phrased in terms of social participation, while the remainder (10 statements) is formulated in terms of social segregation. The key themes ‘friendships/relationships’ and ‘acceptance by classmates’ are represented only by statements phrased in terms of social participation. Table 3 gives some examples of both types of statements.
Table 3. Examples of statements of the Social Participation Questionnaire

<table>
<thead>
<tr>
<th>Key theme</th>
<th>Statement representing social participation</th>
<th>Statement representing social segregation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Friendships/relationships</td>
<td>‘the pupil is a member of a group of friends’</td>
<td></td>
</tr>
<tr>
<td>Contacts/interactions</td>
<td>‘in free time (e.g. during recess) the pupil plays with classmates’</td>
<td>‘classmates regularly exclude the pupil from activities’</td>
</tr>
<tr>
<td>Pupil’s social self-</td>
<td>‘the pupil has the feeling s/he belongs to the group’</td>
<td>‘the pupil feels lonely at school’</td>
</tr>
<tr>
<td>perception</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acceptance by classmates</td>
<td>‘classmates consider the pupil as a full member of the group’</td>
<td></td>
</tr>
</tbody>
</table>

5.2.3 Analysis

Item Response modelling (see e.g. Embretson & Reise, 2000; Hambleton, Swaminathan & Rogers, 1991; Van der Linden & Hambleton, 1997) was used to assess the quality of the individual statements of each of the four subscales, and to examine their mutual separability. An Item Response Model (IRM) expresses the links between item responses and a latent trait, which pertains to the measured concept. For polytomous items, per item each response category is related to the latent trait via an item response function (IRF). An IRF may follow a specific function, yielding parametric IRMs, or meet certain assumptions, yielding nonparametric IRMs. We considered both nonparametric and parametric IRMs (the normal-ogive multidimensional model [McDonald, 1997]), but as the latter failed to fit the data, we only discuss the nonparametric IRMs.

The two nonparametric models considered here are based on three assumptions: unidimensionality (i.e., all items in the subscale measure the same latent trait), local independence (i.e., a pupil’s score on an item is not influenced by the scores on the other items of the scale) and monotonicity (i.e., all IRFs are monotone, non-decreasing). These three assumptions comprise the Monotone Homogeneity Model (MHM; Junker & Sijtsma, 2001; Mokken, 1971; Sijtsma & Molenaar, 2002). When those assumptions are met for a set of items, the important implication is that the individuals’ sum scores of the items provide the ordering of the individuals on the latent trait. When the three assumptions of the MHM model hold, and the IRFs do not intersect, the more restrictive Double Monotonicity Model (DMM) holds. The nonintersection of the IRFs implies
invariant item ordering, which means that the items have the same ordering in terms of probabilities of positive responses to those items (or item category), for all values of the latent trait.

For each of the subscales, the fit of both the MHM and DMM is assessed using the programme Mokken Scale analysis for Polytomous items (MSP; Molenaar & Sijtsma, 2000). For each item, the Item Scalability Coefficient $H_i$ indicates whether item $i$ fits the MHM, because under the MHM $0 \leq H_i \leq 1$. Moreover, it indicates to what extent the item aids in discriminating across individuals. To ensure items meet the MHM assumptions and to have sufficient discrimination power, it is generally accepted that $H_i$ should exceed 0.3 (Mokken, 1971). Coefficient $H$ is a weighted mean of the $H_i$-values of a set of items making up a scale, and indicates the degree to which pupils can be ordered by means of the sum score on the items. With $0.3 \leq H < 0.4$ a subscale is regarded as weak, with $0.4 \leq H \leq 0.5$ as moderate and with $H > 0.5$ as strong. Apart from the Item Scalability Coefficients, other diagnostics are used to assess whether the monotonicity assumption and the non-intersection assumption of the DMM hold for each item. Those diagnostics are summarised into criteria values, where a value $\geq 80$ strongly suggests an assumption violation, values between 40 and 80 are questionable, and a value $\leq 40$ is satisfying (Molenaar & Sijtsma, 2000). To assess possible differential item functioning across the groups of pupils with and without special needs, the ordering of the response categories of all items is compared across groups. A different ordering suggests differential item functioning, which would imply that the scales could not be used to compare individuals from both groups.

To assess the separability of the subscales, the automated item selection procedure (AISP) in MSP will be used, following the guidelines of Sijtsma and Molenaar (2002; pp. 80-82). The AISP aims at finding clusters of items, such that within each cluster the Item Scalability Coefficients exceed a (user-specified) boundary, and between clusters the Item Scalability Coefficients are lower than the boundary. To this end, a stepwise selection procedure is applied. The AISP will be performed on all items using various boundary values, which will indicate whether the items pertain to one scale only or whether they are separable into more than one subscale. When more subscales are distinguished in the AISP, it will be evaluated whether the items are actually assigned to the prespecified subscales.
To sum up, the following analysis strategy will be used: First, for each subscale the separate statements will be assessed using the Scalability Coefficients, criteria values to assess the assumptions of the MHM and the DMM, and the diagnostics for differential item functioning across the two groups. If necessary, and after considering their content, statements may be deleted from the subscales. Second, the scalability of each resulting subscale is assessed. Third, the separability of the subscales will be examined.
When the subscales are in their definite composition, the reliability of both the subscales and the total Questionnaire will estimated using Rho. Rho suffers from bias to a lesser extent than classical reliability estimates (Sijtsma & Molenaar, 2002). According to Nunnally and Bernstein (1994, pp. 265), reliabilities of 0.80 and more are satisfactory in basic research.
Finally, discriminant validity of the entire Questionnaire and the four subscales will be examined by comparing the scores of pupils with special needs with those of typical pupils. The former are expected to receive lower scores on the Questionnaire and on each of the four subscales. As there is no overlap of content between the criteria used to classify the pupils and the Questionnaire’s statements, a confirmation of this expectation implies evidence for the Questionnaire’s discriminant validity.

5.3 Results

5.3.1 Mokken Scale Analysis

The four subscales were analysed separately to identify deviant statements, and to assess possible differential item functioning across pupils with and without special needs. In the following, only the examination of the properties of the 'Friendships/relationships’ subscale will be described in detail. Examination of the other subscales was done similarly. The qualities of the statements of those subscales will be described briefly; a detailed description can be obtained from the first author.

'Friendships/relationships’ subscale. For the six statements of this subscale it was found that $0.40 \leq H_i \leq 0.55$. For each of the statements, the criteria values did not indicate any violation of the monotonicity assumption, but suggested a
violation of the non-intersection assumption for one statement (‘after school, the pupil plays with one or several classmates’). Because this statement showed much overlap with another statement (‘the pupil makes engagements with one or more classmates to play after school’) in terms of content, it was removed from the scale. The scalability of both the remaining five statements and the resulting subscale appear intermediately strong, at $0.42 \leq H_i \leq 0.54$, and $H=0.49$ for the entire subscale.

Comparison of the ordering of the response categories for all statements across the two groups revealed a slight indication of differential item functioning of one statement (‘the pupil is being invited to birthday parties’). Inspection of the order revealed that for pupils with relatively low scores on the Friendship scale, pupils without special needs are invited more often than pupils with special needs, but the difference appeared to be small. On the basis of its content and its sufficient $H_i (0.42)$, it was decided to retain this statement in the subscale. Table 4 presents an overview of the quality of the statements of this subscale.

<table>
<thead>
<tr>
<th>Table 4. $H_i$ Coefficients and criteria values of statements ‘Friendships/relationships’ subscale</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Statement</strong></td>
</tr>
<tr>
<td>Selected statements</td>
</tr>
<tr>
<td>‘the pupil is a member of a group of friends’ (8)</td>
</tr>
<tr>
<td>‘the pupil has one or more friends in the classroom’ (11)</td>
</tr>
<tr>
<td>‘the pupil makes engagements with one or more classmates to play after school’ (19)</td>
</tr>
<tr>
<td>‘one or more classmates invite the pupil to play during holidays’ (25)</td>
</tr>
<tr>
<td>‘the pupil is being invited to birthday parties’ (28)</td>
</tr>
<tr>
<td>Removed statement</td>
</tr>
<tr>
<td>‘after school, the pupil plays with one or more classmates’ (21)</td>
</tr>
</tbody>
</table>

‘Contacts/interactions’ subscale. For 11 out of 14 statements, we found that $0.42 \leq H_i \leq 0.65$, while for the remaining three statements it was revealed that $0.35 \leq H_i \leq 0.39$. It was decided to remove these three statements from the subscale. For the remaining 11 statements, the criteria values did not indicate any violation of the monotonicity assumption. However, for two statements, the criteria values indicated a violation of the non-intersection assumption (relevant criteria values of 70 and 112 for the one statement; relevant criteria values of
Both statements were removed from the subscale. The five removed statements had weak discrimination power, since almost all pupils received a high score on these statements (mean scores range from 4.4 to 4.8). The discrimination power of the remaining nine statements is stronger, as the distribution of scores on these statements is more diverse (mean scores range from 3.5 to 4.1).

The scalability of the nine statements appears strong, $0.62 \leq H_i \leq 0.74$. The same applies to the resulting subscale, $H=0.70$. Differential item functioning appeared to be absent, which implies that the scale scores are directly comparable across pupils with and without special needs.

'Social self-perception of pupil’ subscale. For five out of seven statements it was found that $0.43 \leq H_i \leq 0.53$, while for the remaining two statements $0.30 \leq H_i \leq 0.34$. For one of these two statements, the non-intersection assumption was violated (relevant criteria values of 107 and 104). Both statements aim at the pupil daring to tell or to ask something in the classroom. These statements are somewhat deviating from the other statements in the subscale, which focus on feelings of the pupil (like happiness, loneliness, sense of belonging) in stead of daring. Both statements were removed from the subscale, after which the scalability of the remaining five items slightly increased ($0.47 \leq H_i \leq 0.56$). For none of the remaining five statements any violation of monotonicity assumption or non-intersection assumption was suggested and there is no indication of differential item functioning. Overall scalability coefficient $H$ is 0.51, which implies a strong subscale.

'Acceptance by classmates’ subscale. For five out of seven statements of this subscale it was found that $0.43 \leq H_i \leq 0.53$, while for the other two statements $H_i$ was 0.36 and 0.33. For each of the remaining five statements, the criteria values did not suggest any violation of both the monotonicity assumption and the non-intersection assumption. The scale scores are directly comparable across pupils with and without special needs, as differential item functioning appeared to be absent. The scalability of the individual five statements and the resulting subscale appears strong, with $0.48 \leq H_i \leq 0.61$ and $H=0.54$ for the entire subscale.
To summarise, ten out of 34 statements were removed because of their low $H_i$ value and/or high criteria values. For each of the four resulting subscales, based on 24 statements, no violations of the four assumptions of the DMM were observed. In addition, for each subscale differential item functioning appeared to be absent, implying that the scale scores are directly comparable across pupils with and without special needs.

### 5.3.2 Separability of the subscales

The separability of the four subscales was examined with the automated item selection procedure (AISP). First, the AISP was applied to all 24 statements at different boundary levels. The results indicated a multidimensional item set, as with increasing boundary values four subscales were distinguished. With a boundary value as high as $c=0.60$, the first selected scale contained all nine statements of the ‘Contacts/interactions’ subscale. The subsequent three scales consisted of two statements each, neatly belonging to the ‘Friendships/relationships’, ‘Pupil’s social self-perception’ and ‘Acceptance by classmates’ subscales, respectively. The remaining nine statements were indicated as non-scalable, which is not surprising given the high boundary value. With boundary values $c$ lower than 0.60, the first selected scale consisted of the ‘Contacts/interactions’ statements, mixed with certain statements from the other three subscales. This is not surprising, because the ‘Contacts/interactions’ subscale itself was found to be much stronger ($H=0.70$) than the remaining three ($0.49<H<0.54$). Presumably, statements of the other three subscales not only contain aspects related to their own key theme, but also aspects connected to contacts and interactions.

Repeating the AISP at different boundary levels with all statements except for the nine statements of the ‘Contacts/interactions’ subscale resulted in a clear distinction of the three remaining subscales. This indicates that the statements which initially were attracted to the ‘Contacts/interactions’ subscale also contain unique aspects that belong to the key themes they are representing. Those results suggest that the four subscales are distinguished to a reasonable extent.
5.3.3 Revised version of the Social Participation Questionnaire

The raw scores on the ‘Contacts/interactions’ subscale range from 9 (9x1) to 45 (9x5), the raw scores on the three other subscales from 5 (5x1) to 25 (5x5). In order to facilitate interpretation of the subscale scores, they were linearly rescaled so that the minimum score on all subscales is 0 and the maximum 25. For example, the raw scores on the ‘Acceptance by classmates’ subscale were rescaled by subtracting 5 and then multiplying by 25/20. The total score on the Questionnaire ranges between 0 and 100.

Reliability

The reliability of the final Questionnaire was estimated by Rho. Rho of the total Questionnaire is 0.95, Rho of the subscales varies between 0.80 and 0.95 (see Table 5), meaning a reasonable reliability (Nunnally & Bernstein, 1994).

Table 5. Scalability ($H$) and reliability (Rho) of final subscales

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Number of statements</th>
<th>Subscale $H$</th>
<th>Rho</th>
</tr>
</thead>
<tbody>
<tr>
<td>Friendships/relationships</td>
<td>5</td>
<td>0.49</td>
<td>0.80</td>
</tr>
<tr>
<td>Contacts/interactions</td>
<td>9</td>
<td>0.70</td>
<td>0.95</td>
</tr>
<tr>
<td>Pupil’s social self-perception</td>
<td>5</td>
<td>0.51</td>
<td>0.82</td>
</tr>
<tr>
<td>Acceptance by classmates</td>
<td>5</td>
<td>0.54</td>
<td>0.83</td>
</tr>
</tbody>
</table>

Correlation between subscales

As suggested by the AISP, there appears a clear distinction between the four subscales. In order to examine the relationships between the subscales, Spearman’s Rho was calculated. The correlations between all four subscales appeared to be strong (see Table 6), implying that the scales are closely connected. The correlation between the subscales ‘Friendships/relationships’ and ‘Contacts/interactions’ is strongest. This seems to be understandable, since friends generally spend a lot of time together. As a result, pupils with many friends presumably will have a large amount of interactions. The strong correlations between the other subscales are explainable too. For instance, having friends is associated with enhanced opportunities to exercise behaviours and acquire skills related to social, emotional, and cognitive growth (Ladd, 1990; Newcomb & Bagwell, 1996). Acquiring these skills may facilitate the acceptance
by peers, which subsequently may have a positive influence on the pupil’s social self-perception. Besides, having more mutual friends is related to a more positive social self-concept (Vandell & Hembree, 1994, in Bagwell, 2004). On the other hand, the lack of friends may negatively influence a pupil’s self-perception (Newcomb & Bagwell, 1996; Parker & Asher, 1993). Parker and Seal (1996) revealed that ‘chronically friendless’ children are perceived by their peers as shy, timid, withdrawn and easily angered. These characteristics might prevent peers from interacting with these pupils, which probably negatively influences the self-concept of the pupils and their acceptance by peers.

Although the subscales strongly correlate, the correlations are not perfect. In addition, it turned out that for pupils with special needs, the correlations between the subscales are less strong, compared to pupils without special needs (see Table 6). The differences between these two groups of pupils are significant with regard to the correlations of the subscales ‘Pupil’s social self-perception’ & ‘Friendships/relationships’, ‘Acceptance by classmates’ & ‘Friendships/relationships’ and ‘Contacts/interactions’ & ‘Pupil’s social self-perception’ (p<0.05). The lower connection between subscales for pupils with special needs gives extra cause for distinguishing the four subscales. If only the total score on the Questionnaire would be viewed, valuable differences in the key themes might remain unnoticed. If also the subscale scores are taken into consideration, this is expected to result in a more proper and balanced view of the social participation of the pupil.

Table 6. Correlations between subscales

<table>
<thead>
<tr>
<th>Contacts/interactions</th>
<th>Friendships/relationships</th>
<th>Pupil’s social self-perception</th>
<th>Acceptance by classmates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contacts/interactions</td>
<td>0.77</td>
<td>(0.70, 0.73)</td>
<td></td>
</tr>
<tr>
<td>Pupil’s social self-perception</td>
<td>0.61</td>
<td>(0.42, 0.64)</td>
<td>(0.59, 0.64)</td>
</tr>
<tr>
<td>Acceptance by classmates</td>
<td>0.54</td>
<td>0.60</td>
<td>(0.47, 0.59)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(first number, without brackets, is correlation for all pupils, first number between brackets is correlation for pupils with special needs, second number between brackets is correlation for pupils without special needs)

All correlations are significant at the 0.01 level
5.3.4 Discriminant validity

To assess the Questionnaire’s discriminant validity, a comparison was made between the scores of pupils with special needs and those of their typical classmates. Based on literature (e.g., De Monchy et al., 2004; Pijl, Frostad & Flem, in press; Scheepstra, Nakken & Pijl, 1999), it was expected that the social participation of pupils with special needs would be assessed less positively than that of their classmates. A distinction was also made between different categories of disabilities. Pupils diagnosed as having autistic spectrum disorders and pupils diagnosed as having behavioural disorders were expected to receive the lowest mean scores on the Questionnaire, as research has shown that these pupils find it particularly difficult to build relationships with typical peers and are at risk of becoming isolated in the classroom (De Monchy et al., 2004; Garrison-Harrell, Kamps & Kravits, 1997). Both groups of pupils were expected to receive low scores on each of the four subscales. Pupils diagnosed as having motor disabilities were expected to receive the highest scores on the four subscales, as their type of disability presumably has the least impact on social functioning in the classroom. In addition, motor disabilities are visible and understandable for classmates, which fosters acceptance (Lewis, 1995, in Laws & Kelly, 2005). Comparisons were made between subgroups of pupils with different categories of disabilities for the scores on the total Questionnaire and for the scores on the four subscales.

*Total Questionnaire.* On average, pupils with special needs got a substantially lower score \((M=67.5, \ SD=16.6)\) than their typical counterparts \((M=80.2, \ SD=15.1)\), and this difference appeared significant \((95\% \ CI \ dif = 10.1;15.3)\). This implies that the social participation of pupils with special needs is assessed lower compared to the social participation of their classmates without special needs. As this confirms the expectation that the social participation of pupils with special needs would be assessed less highly, it can be regarded as evidence for the Questionnaire’s discriminant validity.

A one-way analysis of variance (ANOVA) revealed significant group differences between pupils with different categories of disabilities: \(F(4, 225)=4.88, \ p<0.05\) (where the learning disabilities and chronic illness categories were excluded since the group sizes were too small). Pupils diagnosed as having behavioural disorders \((M=62.9, \ SD=18.3)\) and autistic spectrum disorders \((M=64.1, \ SD=15.6)\) got the lowest scores, which confirms our expectation. The
mean score of pupils diagnosed as having intellectual disabilities is slightly higher
\((M=67.6, \ SD=16.0)\). Pupils diagnosed as having speech/language disabilities
\((M=74.3, \ SD=14.1)\) and pupils diagnosed as having motor disabilities \((M=73.2, \ SD=16.5)\) have the highest mean scores. Scheffé post hoc comparison shows
that the social participation of pupils diagnosed as having autistic spectrum
disorders is assessed significantly lower than that of pupils diagnosed as having
speech/language disabilities \((p<0.05)\). There were no significant differences
between other categories of disabilities.

Subscales. Closer inspection revealed that for the four subscales the same
pattern arises as for the total Questionnaire. Corresponding to the scores on the
total Questionnaire, for each of the subscales the mean scores of pupils with
special needs are significantly lower than those of their typical classmates (see
Table 7).

Table 7. Mean scores on subscales of pupils with \((n=233)\) and without special needs
\((n=346)\)

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Mean score, pupils with special needs</th>
<th>Mean score, pupils without special needs</th>
<th>95% CI dif</th>
</tr>
</thead>
<tbody>
<tr>
<td>Friendships/relationships</td>
<td>13.9 ((SD=6.5))</td>
<td>19.1 ((SD=6.5))</td>
<td>4.2;6.2</td>
</tr>
<tr>
<td>Contacts/interactions</td>
<td>18.0 ((SD=4.4))</td>
<td>21.8 ((SD=3.5))</td>
<td>3.2;4.4</td>
</tr>
<tr>
<td>Pupil’s social self-perception</td>
<td>18.7 ((SD=4.9))</td>
<td>21.1 ((SD=4.1))</td>
<td>1.7;3.2</td>
</tr>
<tr>
<td>Acceptance by classmates</td>
<td>16.8 ((SD=4.8))</td>
<td>18.1 ((SD=4.8))</td>
<td>0.5;2.1</td>
</tr>
</tbody>
</table>

An ANOVA on the ‘Acceptance by classmates’ subscale showed no
significant differences between pupils with different categories of disabilities: \(F\)
\((4, \ 227)=1.70, \ p=0.15\). ANOVAs on the ‘Friendships/relationships’,
‘Contacts/interactions’ and ‘Pupil’s social self-perception’ subscales revealed
significant differences between pupils with different categories of disabilities
(respectively, \(F(4, \ 227)=3.50, \ p<0.05\); \(F(4, \ 229)=6.55, \ p<0.05\); \(F(4, \ 228)=2.98, \ p<0.05\)). Scheffé post hoc comparison revealed that the scores on
the ‘Friendships/relationships’ subscale significantly differed between pupils
diagnosed as having speech/language disabilities and pupils diagnosed as having
autistic spectrum disorders, whereby the latter received the lower scores, as
hypothesised. In addition, the scores on the ‘Contacts/interactions’ subscale
significantly differed between pupils diagnosed as having speech/language
disabilities and pupils diagnosed as having motor disabilities on the one hand,
and pupils diagnosed as having autistic spectrum disorders on the other hand, where as hypothesised the latter showed lower scores. For an overview of the results, we refer to Table 8.

Table 8. Mean scores on subscales of pupils with different categories of disabilities

<table>
<thead>
<tr>
<th></th>
<th>Behavioural disorder</th>
<th>Autistic spectrum disorder</th>
<th>Motor disability</th>
<th>Intellectual disability</th>
<th>Speech/language disabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td># pupils (as % total)*</td>
<td>29 (12.2)</td>
<td>97 (40.9)</td>
<td>35 (14.8)</td>
<td>26 (11.0)</td>
<td>47 (19.8)</td>
</tr>
<tr>
<td>Friendships/relationships</td>
<td>13.2 (SD=7.4)</td>
<td>12.6 (SD=5.9)</td>
<td>15.6 (SD=7.0)</td>
<td>13.0 (SD=7.5)</td>
<td>16.3 (SD=5.5)</td>
</tr>
<tr>
<td>Contacts/interactions</td>
<td>17.0 (SD=4.7)</td>
<td>16.9 (SD=4.2)</td>
<td>19.9 (SD=3.9)</td>
<td>17.5 (SD=4.3)</td>
<td>19.9 (SD=3.8)</td>
</tr>
<tr>
<td>Pupil’s social self-perception</td>
<td>17.5 (SD=4.9)</td>
<td>17.7 (SD=5.2)</td>
<td>19.8 (SD=4.7)</td>
<td>19.8 (SD=4.6)</td>
<td>20.0 (SD=3.7)</td>
</tr>
<tr>
<td>Acceptance by classmates</td>
<td>15.1 (SD=5.2)</td>
<td>16.7 (SD=4.6)</td>
<td>17.9 (SD=4.6)</td>
<td>17.3 (SD=4.6)</td>
<td>17.4 (SD=4.9)</td>
</tr>
</tbody>
</table>

*Due to the small number of pupils with learning disabilities (n=2) and pupils with chronic illness (n=1), no data about these pupils are included in the Table.

5.4 Discussion

In this paper, the psychometric qualities of a new teacher questionnaire to assess the social participation of pupils with special needs in regular primary schools were examined. A Mokken Scale Analysis was conducted for this purpose. Based on the results of various analyses, ten out of 34 statements were removed from the Questionnaire.

As for each of the resulting subscales the Mokken’s double monotonicity model appeared to fit well, the subscale scores are on an ordinal scale, and the separate statements per subscale are invariantly ordered along the subscale. The scale scores are directly comparable across pupils with and without special needs, because differential item functioning appeared to be absent. Subsequent analyses did support the division of social participation into the four distinct key themes. The Questionnaire as a whole and its subscales turned out to be sufficiently reliable. Evidence for the Questionnaire’s discriminant validity was also obtained as the expectation that the social participation of pupils with special needs would be assessed as less high using the Questionnaire could be confirmed: both on the total Questionnaire and on each of the four subscales,
the mean scores of pupils with special needs were significantly lower than those of their typical counterparts. More aspects of the Questionnaire’s construct validity will be examined in further research.

Scheffé post hoc comparisons of the scores on the Social Participation Questionnaire made clear that there are significant differences between pupils with different categories of disabilities. It was revealed that the scores on two out of four subscales significantly differed between pupils diagnosed as having autistic spectrum disorders and pupils diagnosed as having speech/language disorders (and for one subscale also pupils diagnosed as having motor disabilities), whereby the latter received the highest scores. The other differences in subscale scores between pupils with different categories of disabilities were not significant. This might be related to the fact that the subgroups of disabilities contained rather small numbers of pupils. We could have decided to involve less categories of disabilities in the study, in order to compose larger subgroups of specific categories of disabilities. However, we consciously included such a variety of disabilities in the sample, as the Questionnaire has been developed for all pupils. This fits with the fact that in the Netherlands, the group of pupils with special needs attending regular education is rather heterogeneous. In order to examine the applicability of the Questionnaire, it was desirable to involve pupils with various categories of disabilities in the study. In a future study, when the Questionnaire has its definitive form, it would be advisable to involve larger homogeneous subgroups of pupils. That way it might be possible to aim analyses on possible differences in scores on the Questionnaire of pupils with different categories of disabilities. Among other things, one could assess possible differential item functioning between subgroups of pupils with different categories of disabilities. In the current study, we do not know if the scale scores are directly comparable across pupils with different categories of disabilities, as the subgroups were too small to conduct analyses. However, on theoretical grounds, we assume that the Questionnaire is applicable for pupils with various categories of disabilities, as none of its statements are particularly focused on specific categories of disabilities.

The low mean scores of the pupils diagnosed as having autistic spectrum disorders (whether or not in combination with behavioural disorders) were in accordance with other studies, which revealed that in particular these pupils experience difficulty building relationships with typical peers and are at risk of
becoming isolated in the classroom (De Monchy et al., 2004; Garrison-Harrell et al., 1997). Further research into differences between categories of disabilities is recommended, since treating pupils with special needs as one homogenous group might reveal a too negative picture of the social participation of pupils with specific types of disabilities.

In conclusion, the Social Participation Questionnaire has the potential to become a useful tool for teachers, as it helps them obtain clarity about the social participation of their pupils. The Questionnaire takes into consideration the comprehensiveness of the concept of social participation by providing insight into four distinguished themes of social participation, while other instruments only focus on parts of this concept (for instance sociometric questionnaires mainly focus on pupils’ friendships and/or their acceptance by classmates, and observation scales mainly focus on interactions). The Social Participation Questionnaire is a teacher-friendly instrument because it can be completed in a short time and provides a reliable and wide picture of the social participation of pupils. Further research into the Questionnaire’s psychometric properties is needed. The outcomes of the present study provided evidence for the Questionnaire’s discriminating power, because of the presence of the expected score differences between pupils with and without special needs. It should be kept in mind that this difference in scores is a necessary, but not sufficient prerequisite for establishing the discriminant validity of our Questionnaire. More aspects of the Questionnaire’s validity will be examined in further research.

After having obtained evidence for the validity of the Social Participation Questionnaire, its usability in practice will be examined. The Questionnaire is expected to help teachers make more appropriate assessments of the social participation of their pupils. As a result, teachers might be encouraged to take measures in time, which will contribute to optimise the situation of pupils with special needs in inclusive classrooms.

5.5 References


