**RESEARCH**

# Assessing social competence and antisocial behaviors in children: item response theory analysis of the home and community social behavior scales

Silje Sommer Hukkelberg[1*] and Björn Andersson[2]

## Abstract

**Background** The Home and Community Social Behavior Scales (HCSBS) is a rating scale that assesses social competence and antisocial behavior among children and youths between ages 5–18. The present study aimed to investigate the psychometric properties of the HCSBS by applying item response theory (IRT).

**Methods** The HCSBS was completed by parents of 551 Norwegian children refereed to three independent interventions towards problem behaviors. Data used in this study was collected before the interventions started. IRT was carried out in R version 4.0.0 to investigate HCSBS items, subscales and main scales.

**Results** The results showed that the two-dimensional IRT models for social competence and antisocial behavior were the most appropriate. The measurement precision of the scales was high for a large range of the latent spectrum, and estimated reliabilities were satisfactory. Model evaluations indicated that the overall model fit for the scales were acceptable, but some misfit existed with respect to specific item pairs.

**Conclusion** The results indicate that the HCSBS is a reliable measurement instrument although there is still a potential for improvement by revising some of the items.

**Keywords** Children, Antisocial behavior, Social competence, HCSBS, Item response theory, Psychometrics

## Background

Children and adolescents with behavioral problems show an increased risk for developing negative outcomes like school dropout, social exclusion, law breaking, psychopathology, and other health-related problems [1–3]. The costs are considerable, for each individual and for

the society at large [4, 5]. Empirical findings show that child problem behaviors often operate in concert with poor social competence [6, 7]. This suggests that it may be advantageous to assess both concepts when children show high levels of resistant problem behaviors and are considered for intervention programs. However, in order to identify, monitor and treat behavioral problems and interpersonal challenges, reliable and validated screening instruments are required.

### Problem behavior and social competence

A growing body of research has demonstrated an inverse association between child conduct problems and social competence [8–11]. That is, children who display high

*Correspondence:
Silje Sommer Hukkelberg
s.s.hukkelberg@nubu.no
[1] The Norwegian Center for Child Behavioral Development (NUBU), Essendropsgt 3, Majorstuen, Postbox 7053, 0306 Oslo, Norway
[2] Centre for Educational Measurement (CEMO), University of Oslo, Blindern, Postboks 1161, 0318 Oslo, Norway

levels of conduct problems tend to show low levels of social competence, or the other way round. Persistent child problem behaviors is a well-known risk factor associated with a troublesome future pathway that may follow into adolescents and sometimes into adulthood [12]. On the other hand, social competence is considered a protective factor associated with positive development and healthy functioning [13, 14] that e.g., facilitate school success [15–17]. Whereas child problem behavior is apparent through aggressive, oppositional or inattentive behaviors that may cause considerable challenges for parents, social competence is characterized by e.g., awareness about others needs and situation- appropriate behaviors [18]. Children with poor social competence can show inadequate social information processing and poor adaption to situations, which further result in rejection and dislike from peers and exclusion from social groups. Nevertheless, both problem behavior and social competence should be considered in relation to a child's age, as both concepts are characteristically dynamic and complex in nature. Some have advocated that the two constructs should be considered separate albeit related dimensions of social functioning [19].

A considerable number of studies show that there is a significant association between child problem behavior and social competence, however, not all children with problem behavior are socially unskilled, or vice versa. It may, in fact, be difficult to disentangle the temporal precedence of the two concepts. Some findings suggest that poor social competence may contribute to maintain or increase the intensity of child problem behaviors [20], as it fosters frustration and limits a child's possibilities to develop positive interactions with others. In addition, it should be noted that the strength between problem behavior and social competence varies considerable across studies, and depend on whether the two concepts are assessed by the same respondents and the bandwidth of the constructs being measured. A recent meta-study based on 54 independent studies among children aged 3–13, showed an overall negative and significant correlation of medium effect size ($r = -0.42$) between behavioral problem and social competence [21]. The magnitude was even higher ($r = -0.47$, $p < 0.001$) when both constructs were assessed by parents, but considerable lower when one of the constructs were reported by a parent and the other by a teacher ($r = -0.17$, $p < 0.001$). However, several of the studies relied on different measurement instruments to assess child problem behavior and social competence, and some date back to the 1980s and early 1990s, and have shown varying psychometric properties [22]. Overall, these findings suggest that it could be worthwhile to consider using instruments that are newer and more psychometrically sound, and that

address both social competence and problem behavior. This would probably ease use and interpretation of both concepts.

## The home and community social behavior scales

The Home and Community Social Behavior Scales (HCSBS [23]) is a parent-reported instrument that provides insight into children's Social Competence (Scale A) and Antisocial Behavior (Scale B) using the same rating scale. The HCSBS scale is the home version of the School Social Behavior Scales (SSBS [24]). Whereas Scale A assesses the big picture of adaptive and positive social behaviors, Scale B measures the broad concept of socially related problem behaviors. Both scales comprise two subscales. The social competence scale measures Peer relations (e.g., assesses positive skills with peers) and Self-management/compliance (e.g., responses to adult expectations), whereas the Antisocial behavior scale includes the Antisocial/aggressive scale (e.g., overt violation, intimidation or harm to others) and the Disruptive/demanding scale (e.g., behaviors likely to disrupt ongoing activities and place inappropriate demands on others). Validation studies indicate that the HCSBS has good psychometric properties [25–27], in addition, two systematic reviews evaluating the Scale A have encouraged the use of the HCSBS to assess social competence [28, 29]. However, a drawback pertaining these studies is that they rely on classical test theory that e.g., treats the ordinal item scores as continuous, and puts emphasis on total scores where all items are equally weighted [30]. Although some studies have considered the social competence scale[31, 32], few studies have evaluated the dimensionality of the HCSBS or its parallel teacher form (SSBS). In the present study we aimed to extend previous knowledge about the HCSBS by applying item response theory (IRT), which is the preferred method for psychometric evaluations [33, 34].

## Item response theory

Item response theory (IRT) comprises a family of flexible statistical models and techniques that can be used to evaluate the psychometric properties of a scale, improve scoring accuracy, increase measurement precision, and form the basis for evaluating item properties [35, 36]. In this study, we use polytomous IRT [37, 38] to model the response options for the ordinal observed variables of the HCSBS. Like classical test theory, IRT assumes that the latent construct cannot be measured directly, but instead is measured by a set of indicators. IRT, however, specifies a non-linear relationship between the continuous latent variable and the categorical observed variables. IRT models are typically characterized by the two parameters discrimination parameters and location

parameters [39]. The discrimination parameters indicate how well the items can differentiate between participants with different values of the latent variable (similar to an index of item discrimination), while the location parameters inform about the average value of the latent variable associated with a particular response category (similar to item difficulty). With basis in the estimated item parameters, IRT allows for a detailed description of item and scale properties.

### The present study

In this study, we utilized IRT to investigate three research questions. First, we investigate dimensionality of the Social competence (Scale A) and the Antisocial behavior (Scale B) scales as measured by HCSBS, by comparing one-dimensional models with two-dimensional models for each scale. For the two-dimensional models, the different subscales of Scale A and Scale B were considered as measuring separate but correlated dimensions of social competence and antisocial behavior, respectively. We evaluate the estimated models with respect to the model fit and item fit. Second, we infer the item and scale properties for the subscales Peer relations (Scale A1) and Self-management/compliance (Scale A2), and Antisocial/aggressive behavior (Scale B1) and Defiant/disruptive behavior (Scale B2), using item category probability curves, item information functions, and test information functions. Third, we estimate IRT model-based reliability indices for sum scores and IRT scores to evaluate the precision of measurement of the scale scores.

The benefits of IRT are a more realistic model for item responses (which accounts for the ordinal nature of the item scores), improved evaluation of item properties, increased measurement precision through pattern scoring, full-information methods for estimation, and access to detailed tools to evaluate scale properties such as scale difficulty via test characteristic curves and scale precision from test information functions. To our knowledge, this is the first study to evaluate the HCSBS using IRT and as such the study provides important additional information on the properties of the HCSBS, which is useful for practitioners and researchers alike.

## Method

### Participants

Data derived from 551 children who were enrolled into three independent interventions aimed to reduce child conduct problems [40–42]. The present data was collected before the interventions started. A parent or a primary caretaker completed the HCSBS questionnaire [23]. The eligible families agreed to participate in the study by signing a written informed consent document. In line with standard procedures in the Norwegian children's

services, the interventions were offered based on practitioners' clinical judgments rather than formal screening. In all samples, children were excluded from participation if they were diagnosed with autism, had been exposed to sexual assaults, or were intellectually disabled, or if parents had serious mental health problems. The Regional Committee for medical research ethics approved all three studies (REK 2.2006.2066).

Sample 1 consisted of children from 137 families, recruited from 11 agencies situated in different municipalities in Norway. Children were between the ages of 3 and 12 ($M = 8.56$, $SD = 2.35$) and 50 (36.5%) were girls. Parents were on average 37.42 years ($SD = 6.34$). Among the participating children, 66 (48.2%) lived with both biological parents, 21 (15.3%), and 50 (36.5%) lived with single parents. The average gross annual family income was 509,610 Norwegian Kroner ($SD = 347.70$), which is approximately $51,873, representing an upper middle income level. According to parent self-report, 37 (27%) had a college or higher university degree, 83 (60.6%) had finished high school, and 17 (12.4%) had completed junior high school or elementary school. Most of the parents had Norwegian background (126 or 92%), only one (0.7%) was from another western European country, and 8 (7.3%) reported "other" ethnicity. Sample 2 included children from 198 families recruited from 9 municipalities in Norway. Child age ranged from 3 to 12 years ($M = 7.64$, $SD = 2.19$), and 39 (19.7%) were girls. The average age of the reporting parent was 36.30 years ($SD = 6.07$). Among the participating children, 106 (53.5%) lived with both biological parents, 29 (14.6%) with a parent cohabiting with another adult, and 63 (31.8%) lived with single parents. The average family income was 564.090 Norwegian Kroner ($SD = 267.05$), which is approximately $ 57,469 and representing an upper middle income level. Parent self-report showed that 80 (40%) had a college or higher university degree, 96 (48.5%) had finished high school, and 22 (11.1%) had completed junior high school or elementary school. Most parents reported a Norwegian background (182 or 92%), and 16 (7.2%) reported "other" ethnicity.

Sample 3 included children from 216 families from five municipalities in Norway. One child was excluded because of an autism diagnosis. Child age ranged from 3 to 12 years ($M = 7.28$, $SD = 2.61$), and 69 (31.9%) were girls. The average age of the parents was 35.31 years ($SD = 6.08$). Hundred and ten children (50.9%) lived with both biological parents, 27 (12.5%) with a parent cohabiting with another adult, and 79 (36.6%) lived with single parents. The average gross annual family income was 539.110 Norwegian Kroner ($SD = 328.29$), which is approximately $54, 827 and represents an upper middle income level. Among the parents, 85 (39.4%) had a

college or higher university degree, 114 (52.8%) had finished high school, and 17 (7.9%) had completed junior high school or elementary school. Most of the parents had a Norwegian background (202 or 93.5%), four (1.9%) were from other western European countries, and the remaining 10 (4.6%) reported "other" ethnicity.

### Instrument
The Home and Community Social Behavior Scales (HCSBS) [23] is a screening instrument that assesses social competence (Scale A) and antisocial behavior (Scale B), which each comprises 32 items rated on a 5-point scale from 0 (never) to 4 (frequently). Scale A consists of the subscales Peer relations (A1; 17 items) and self-management/compliance (A2; 15 items) and includes evaluations of items like "completes chores without being reminded", "shows self-control", and "is invited by peers to join in activities". Higher scores indicate higher levels of social competence. Scale B measures antisocial behavior and contains the two dimensions Antisocial/aggressive (B1; 17 items) and defiant/disruptive (B2; 15 items) behaviors. Items include e.g., "gets into fights", "bothers and annoys others", and "acts impulsively without thinking". Higher scores on this scale indicate more antisocial behavior. The HCSBS was translated from English to Norwegian by the Norwegian Center for Child Behavioral Development.

### Statistical analysis
We used the statistical programming environment R Version 4.0.0 [43] and the package mirt [44] to estimate the IRT models and evaluate item and scale properties.

*Item response theory (IRT) analysis* We estimated unidimensional and multidimensional graded response models (GRM) [38] and generalized partial credit models (GPCM) [37] for the ordinal observed variables using marginal maximum likelihood estimation [45]. The functional form of the IRT model was decided based on the AIC [46]. To characterize the properties of the items on the scales, we examined item category probability curves and item information functions [47]. To assess the measurement precision of the scales for different values of the latent variable, we computed scale information functions [48].

*Dimensionality* We evaluated the dimensionality of each of the scales (A and B) by fitting a series of models. First, we specified unidimensional models for each of the subscales Peer relations (17 items, Scale A1), self-management/compliance (15 items, Scale A2), Antisocial/aggressive (17 items, Scale B1) and defiant/disruptive (15 items, Scale B2). If the unidimensional models for each

subscale were considered acceptable, we specified models for all the items in Scales A and B, respectively. Hence, for the items in Scale A, we specified (1) a unidimensional model and (2) a two-dimensional model where the correlated latent variables were related to the Peer relations or self-management/compliance items. For the items in Scale B, we specified (1) a unidimensional model and (2) a two-dimensional model where the correlated latent variables were related to Antisocial/aggressive or defiant/disruptive items. For either scale, we compared the model fit between the unidimensional and two-dimensional models using likelihood ratio tests.

*Model evaluation* Model fit was assessed via the limited-information fit statistic M2 [49], with the fit statistics root mean square error of approximation ($RMSEA_2$) and the standardized root mean square residuals (SRMSR), where $RMSEA_2 \leq 0.089$ and $SRMSR < 0.05$ indicated good approximate model fit [50], and with residual analysis [51]. Residual correlations larger than 0.1 in absolute value were further investigated with respect to the content of the corresponding pairs of items. Fit of individual items was evaluated using $S - \chi^2$ item fit statistics [52, 53].

*Reliability estimation* The reliability of the sum scores were estimated using Monte-Carlo computer simulations [54]. We also estimated the reliability of expected á posteriori (EAP) IRT scores with the approach defined in Kim [55].

## Results
### Descriptive statistics
The mean of Scale A1 was 41.52 (SD = 12.16, range: 4–68) whereas Scale A2 had mean 31.96 (SD = 9.81, range: 7–59). The mean of Scale B1 was 22.04 (SD = 10.66, range: 0–56) and Scale B2 had mean 22.37 (SD = 10.86, range: 0–57). The total scales showed means of 73.21 (SD = 20.14, range 16–127) and 44.41 (SD = 20.8, range 3–108), for Scale A and B, respectively. The total scores correlated $r = -0.59$ ($p < 0.001$). Additional file 1: Tables S1 and S2 provide item means (SD) and item-total correlations for Scales A and B. By convention, inter-item correlations equal or below 0.4 indicate very good discrimination. Only one item of Scale A showed poor discrimination (item 17) while three items were below 0.4 for Scale B (items 2, 13, and 32). Based on the content of the items, we still included them in the subsequent analysis.

## IRT model selection and model evaluation

The model selection resulted in the GRM being preferred over the GPCM for all subscales. A series of IRT models were then estimated to investigate the assumptions of one-dimensional scales, and an overarching one- or two-dimensional model within social competence and antisocial behavior, respectively. These results are presented in Table 1, showing that scales A1 and B2 had good approximate fit according to the $RMSEA_2$ (0.080 and $0.071 \leq 0.089$) but some lack of fit according to the SRMSR (0.069 and $0.082 > 0.05$). Meanwhile, A2 and B1 showed evidence of some misfit according to both the $RMSEA_2$ (0.109 and $0.108 > 0.089$) and the SRMSR (0.071 and $0.074 > 0.05$). However, we judged the fit of the unidimensional models for each subscale to be sufficiently good to pursue further analyses. Regarding the models for the total scales of social competence (scale A) and antisocial behavior (scale B), scale A showed significantly better model fit as a two-dimensional model rather than a one-dimensional model ($\chi^2(df)$: 609.172 (1), $p < 0.001$). The correlation between the two latent variables representing Peer relations and Self-management/compliance was estimated to be 0.725 (95% CI 0.675, 0.775). The model fit for the two-dimensional model showed good approximate fit with respect to the $RMSEA_2$ ($0.073 \leq 0.089$) but some misfit considering the SRMSR ($0.077 > 0.05$). For scale B, the two-dimensional model also fitted better than the one-dimensional model ($\chi^2(df)$: 6.869 (1), $p = 0.009$) but the correlation between the latent variables representing Antisocial/aggressive and Defiant/disruptive was estimated to be very high at 0.979, with 95% CI (0.962, 0.995). The model fit of the two-dimensional model showed good approximate fit with respect to the $RMSEA_2$ ($0.088 \leq 0.089$) but some misfit with the SRMSR ($0.079 > 0.05$). The item fit statistics are included in Additional file 1: Tables S3 and S4. Two items on each scale had significant misfit at a significance level of 0.05: items 4 ($S - \chi^2(df) = 131.5$ (106); $p = 0.047$), and 9 ($S - \chi^2(df) = 161.1$ (128); $p = 0.025$) of Scale A and items 9 ($S - \chi^2(df) = 149.1$ (122); $p = 0.048$) and 25 ($S - \chi^2(df) = 129.013$ (103); $p = 0.042$) of Scale B. After adjusting the significance level with the Bonferroni-Holm correction [56], none of the items displayed a lack of fit and we concluded that the item fit was acceptable.

The estimated reliabilities for each scale are presented in Table 2. The reliability of the sum scores for A1 was estimated to be 0.920 and the reliability for IRT scores 0.928, whereas the sum scores for A2 was estimated to be 0.894 and the reliability for IRT scores 0.912. Considering antisocial behaviors, the estimated reliability for B1 was 0.887, and for the IRT scores 0.939. B2 had estimated sum score reliability 0.902 and IRT score reliability 0.942.

**Table 1** Model fit statistics for unidimensional and two-dimensional models

| Scale/models | BIC | $M_2$ (df) | *p* | RMSEA | (95% CI) | SRMSR |
|---|---|---|---|---|---|---|
| Peer relations (Scale A1, 1D) | 24,186.692 | 304.54 (68) | < 0.001 | 0.080 | (0.069; 0.090) | 0.069 |
| Self-manag/compliance (Scale A2, 1D) | 21,447.134 | 337.62 (45) | < 0.001 | 0.109 | (0.096; 0.122) | 0.071 |
| Antisocial/aggressive (Scale B1, 1D) | 23,069.409 | 502.651 (68) | < 0.001 | 0.108 | (0.097; 0.118) | 0.074 |
| Defiant/disruptive (B2, 1D) | 21,345.987 | 171.017 (45) | < 0.001 | 0.071 | (0.058; 0.085) | 0.082 |
| Social competence (1D) | 45,941.976 | 1624.549 (368) | < 0.001 | 0.079 | (0.075; 0.083) | 0.085 |
| Antisocial behavior (1D) | 43,640.189 | 1944.889 (368) | < 0.001 | 0.088 | (0.084; 0.092) | 0.079 |
| Social competence (2D) | 45,339.114 | 1453.484 (367) | < 0.001 | 0.073 | (0.069; 0.078) | 0.077 |
| Antisocial behavior (2D) | 43,639.63 | 1932.957 (367) | < 0.001 | 0.088 | (0.083; 0.093) | 0.079 |

Scale A assesses social competence (A1: peer relations, A2: self-management/compliance), scale B assesses antisocial behavior (B1: antisocial/aggressive, B2: defiant/disruptive). In IRT, $M_2$ (Maydeu-Olivares and Joe 2005) estimates global model-data fit

1D, one-dimensional; 2D, two-dimensional; BIC, Bayesian Information Criterion; df, degrees of freedom; RMSEA, root mean square error of approximation; SRMSR, standardized root mean square residual

**Table 2** Reliability coefficient estimates

| | Social competence (Scale A) | | | Antisocial behavior (Scale B) | | |
|---|---|---|---|---|---|---|
| | **Total scale** | **Peer relations** | **Self-management/ compliance** | **Total scale** | **Antisocial/ aggressive** | **Defiant/ disruptive** |
| Sum scores | 0.945 | 0.920 | 0.894 | 0.944 | 0.887 | 0.902 |
| IRT EAP scores | – | 0.928 | 0.912 | – | 0.939 | 0.942 |

EAP, expected á posteriori; IRT, item response theory

Overall, the results indicate that both scales are highly reliable.

## Item and scale properties

Figure 1 shows the expected scale score functions for the subscales Social competence and Antisocial behavior, respectively. The results show that the scales are well-matched relative to the population under study, with very few respondents expected to score at the minimum or maximum of the respective scales.

To illustrate the measurement precision of the scale scores based on the fitted two-dimensional model, we plotted the scale information functions for each of the subscales (Fig. 2). Scale information functions are the sum of the item information functions and indicate the precision of the scale as a whole for different values of the underlying latent variable [57]. Generally speaking, the scale information functions indicate that both scales

have a high measurement precision across the latent spectrum, with peaks around or higher than 10 and lows that are above 5 for the most relevant ranges of the latent spectrum. The social competence scales have peaks at the middle range of the latent spectrum, and both scales provide the highest measurement precision for values of the latent variable between −2 and 2, which is expected to cover 95% of the respondents. On the other hand, both of the subscales of antisocial behavior have somewhat higher measurement precision for higher values of the latent variables, and the Antisocial/aggressive scale provides the most information between −1 and 2, whereas the Defiant/disruptive scale provides the most information between 0 and 3. Much of the same pattern occurred at the item-level (Additional file 1: Figs. S3 and S4), as illustrated in Fig. 3 by item 6 ("understands problems", scale A1) and item 27 ("boasts", scale B2).
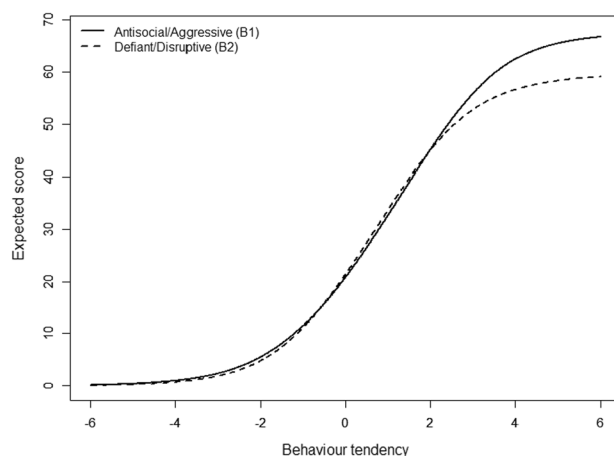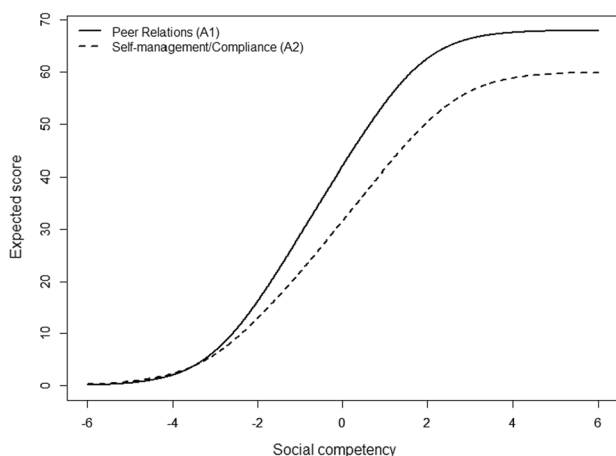


**Fig. 1** Expected scale score functions
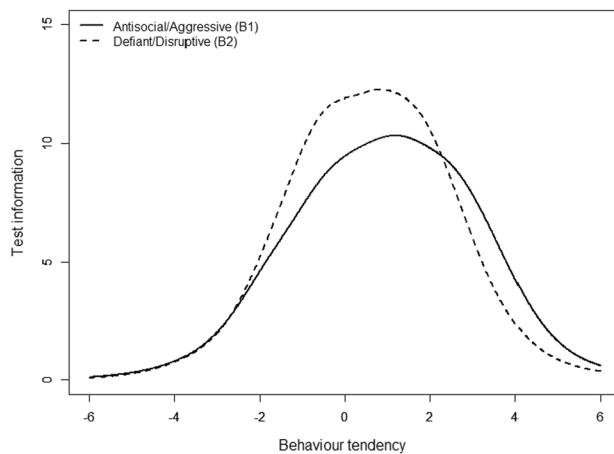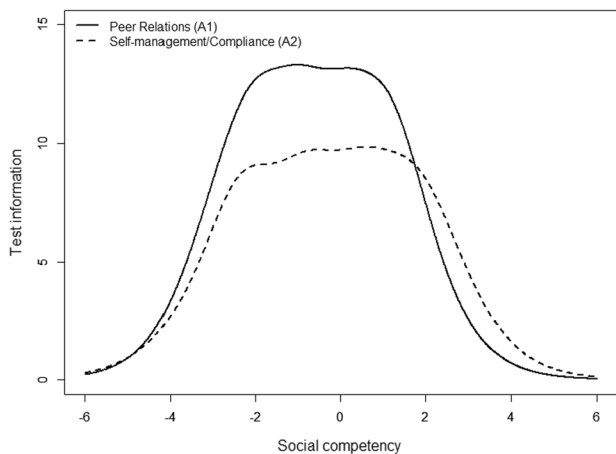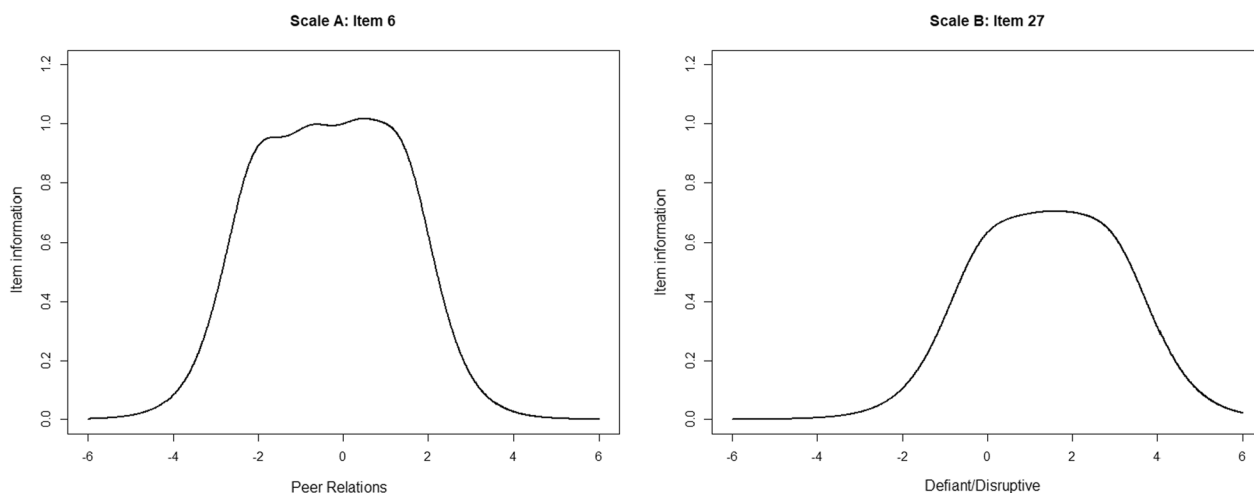


**Fig. 2** Scale information functions

**Fig. 3** Information functions for item 6 of scale A and item 27 of scale B

### Residual analysis

We conducted detailed analysis of the residual correlation matrix to identify sources of misfit in the model. Table 3 shows the highest residual correlations within each of the four scales. Residual correlations for the Peer relation scale involves two items that show overlap in content (empathy; items 6 and 22), whereas in the Self-management/compliance scale we found two items with the same wording (completes chores; item 13 and 14). Considering the Antisocial/aggressive scale, substantial residual correlations were found for item 5 and item 25 that both reflect aggressive behavior, and item 6 and 28 that reflect dishonesty. Lastly, the residual correlations in the Defiant/disruptive scale reflect overlap in content (headstrong; item 2 and 13) and highlighting oneself (item 3 and 20, and item 7 and 27).

**Table 3** The highest residual correlations within each of the scales

|  | Residual correlation |
|---|---|
| *A1: Peer relations* | |
| #22 sensitive to feelings * #30 invited by friends | 0.15 |
| #6 understands problems * #30 invited by friends | 0.14 |
| #6 understands problems * #22 sensitive to feelings | 0.14 |
| *A2: Self-management/compliance* | |
| #13 completes tasks * #14 completes tasks on time | 0.20 |
| #17 behaves at school * #18 asks for help | 0.19 |
| #23 responds appropriately * #24 controls temper | 0.17 |
| #24 controls temper * #31 shows self-control | 0.17 |
| *B1: Antisocial/aggressive scale* | |
| #5 fights * #25 trouble | 0.26 |
| #6 is dishonest * #28 not trustworthy | 0.19 |
| #5 fights * #28 not trustworthy | 0.17 |
| *B2: Defiant/disruptive* | |
| #3 defiant * #20 insults friends | 0.21 |
| #2 takes things * #13 not share | 0.21 |
| #7 teases * #23 difficult to control | 0.21 |
|  | 0.19 |
| #7 teases * #27 boasts | 0.18 |

## Discussion

In the present study, we applied IRT to investigate the psychometric properties of the social competence and Antisocial behavior scales in the HCSBS. To our knowledge this is the first study to use IRT to examine the HCSBS. Both social competence and antisocial behavior have received extensive attention in the literature on child development and adjustment, but to advance knowledge about these constructs, valid and reliable instruments are required. Our results showed acceptable model-fit for the two-dimensional structure of social competence, and Peer relations and Self-management/compliance appear to measure different but related aspects of social competence. Considering the Antisocial behavior scale, results showed a strong correlation between the Antisocial/aggressive and Defiant/disruptive scales, suggesting that the two concepts can not empirically be considered distinct from each other. Reliabilities for Social competence and Antisocial behavior subscales and total scale-scores were high. Furthermore, our analysis revealed that individual items varied in terms of how much information they contributed to the overall measurement precision of the scale. The scales also had different properties in terms of measurement precision. Peer relations showed higher test information across the latent spectrum than the Self-management/compliance scale, whereas the Defiant/disruptive scale had higher overall test information compared to the Antisocial/aggressive scale. All scales showed most information around the middle of the latent spectrum and can thus be viewed as well-matched with the intended target population. Inspections of residual correlations indicated that several items, in both scales, should be further evaluated and potentially revised in future versions of the scale.

The study has several implications for research and practice. First, the HCSBS is one of few scales that measures problem behavior and social comptence using the same rating scale. This makes it easier to use and interpret the relation between the two concepts. Our findings from a sub-clinical sample of Norwegian children indicate that interpretations based on the subscale scores of the HCSBS is supported, and that the sum scores and IRT scores of the subscales are highly reliable. However, practitioners should be aware that the HCSBS does not distinguish well between Antisocial/aggressive and Defiant/disruptive behaviors. Our item information analyses showed that the social competence and antisocial behaviors scales were most informative around the mean of the latent spectrum. As a result, the scales are well-matched with the target population, but do not very precisely assess respondents with very low or very high values. Thus, it seems most appropriate to measure antisocial behavior based on the combined scores from both subscales. The two scales measuring social competence do however provide distinct and separate scoring. It should be noted, that although there is broad consensus that social competence comprises interpersonal behaviors associated with a positive social outcome [58, 59], the indicators vary considerably [60, 61]. Last, it is unknown whether the combination of poor social competence and antisocial behavior is predictive of more serious child problems compared to other factors. This should be examined further. Several limitations apply to this study. First, our sample consisted of a low number of girls ($n = 157$), which made it impossible to consider gender differences. Findings indicate gender differences in both social competence and antisocial behavior [13, 62], thus such analyses would provide valuable information. Furthermore, our results are based on a sub-clinical sample of children, and do not necessarily generalize to other samples. It is also unclear if the misfit we identified for some items relates to the Norwegian version only, so more studies on other translations are warranted. Modifications of specific items would however improve the instrument further.

## Conclusion

The findings of the present study add new knowledge about the psychometric properties of HCSBS. Our results indicate that the HCSBS has clinical utility when it comes to measure social competence and antisocial behavior in children. Thus, practitioners can rely on one instrument only instead of using different instruments to measure and monitor social competence and antisocial behavior in children with social and behavioral difficulties.

## Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40359-023-01045-1.

---

**Additional file 1: Table S1**. Descriptive statistics for the social competence (Scale A). **Table S2**. Descriptive statistics for the antisocial scale (Scale B). **Table S3**. Item fit statistics for scale A. **Table S4**. Item fit statistics for scale B. **Figure S1**. Item category response functions for Scale A. **Figure S2**. Item category response functions for scale B. **Figure S3**. Item information functions for scale A. **Figure S4**. Item information functions for Scale B.

---

## Author's information

**Silje Sommer Hukkelberg** is a senior researcher at The Norwegian Center for Child Behavioral Development (NUBU).
**Björn Andersson** is associate professor at the Centre for Educational Measurement (CEMO).

## Availability of data and materials
The data that support the findings of this study are not publicly available as they are part of a because of the sensitivity of the information. Data are however available from the local research committee (post@nubu.no) or the corresponding author upon reasonable request.

## Declarations

### Ethics approval and consent to participate
The study was approved by the Regional Committee for Medical and Health Research Ethics and the Norwegian Social Sciences Data Services (REK 2.2006.2066). All procedures are in accordance with the ethical standards of the institution, the research committee and the Declaration of Helsinki. Written informed consent was obtained from all parents/or legal guardian.

### Consent for publication
Not applicable.

### Competing interests
The authors declare that they have no competing interests.

## References
1. Coie JD, Dodge KA. Aggression and antisocial behavior. In: Damon W, Eisenberg N, editors. Handbook of child psychology: social, emotional, and personality development, vol. 3. Hoboken: Wiley; 1998. p. 779–862.
2. Wertz J, Agnew-Blais J, Caspi A, Danese A, Fisher HL, Goldman-Mellor S, Moffitt TE, Arseneault L. From childhood conduct problems to poor functioning at age 18 years: examining explanations in a longitudinal cohort study. J Am Acad Child Adolesc Psychiatry. 2018;57(1):54–60. https://doi.org/10.1016/j.jaac.2017.09.437.
3. Loeber R, Farrington DP. Child delinquents: development, intervention, and service needs. Thousand Oaks: Sage; 2001.
4. Foster EM, Jones DE, Conduct Problems Prevention Research Group. The high costs of aggression: public expenditures resulting from conduct disorder. Am J Public Health. 2005;95(10):1767–72. https://doi.org/10.2105/AJPH.2004.061424.
5. Nystrand C, Feldman I, Enebrink P, Sampaio F. Cost-effectiveness analysis of parenting interventions for the prevention of behavior problems in children. PLoS ONE. 2019;14(12):e0225503. https://doi.org/10.1371/journal.pone.0225503.
6. Chen SH, Hua M, Zhou Q, Tao A, Lee EH, Ly J, Main A. Parent–child cultural orientations and child adjustment in Chinese American immigrant families. Dev Psychol. 2014;50(1):189. https://doi.org/10.1037/a0032473.
7. Montroy JJ, Bowles RP, Skibbe LE, Foster TD. Social skills and problem behaviors as mediators of the relationship between behavioral self-regulation and academic achievement. Early Child Res Q. 2014;29(3):298–309.
8. Hukkelberg S, Ogden T, Tømmerås T. Parent training: effects beyond conduct problems. Child Youth Serv Rev. 2018. https://doi.org/10.1016/j.childyouth.2019.03.009.
9. Hosokawa R, Katsura T. A longitudinal study of socioeconomic status, family processes, and child adjustment from preschool until early elementary school: the role of social competence. Child Adolesc Psychiatry Ment Health. 2017;11(1):1–28. https://doi.org/10.1186/s13034-017-0206-z.
10. Sigmarsdóttir M, Thorlacius Ö, Guðmundsdóttir EV, DeGarmo DS. Treatment effectiveness of PMTO for children's behavior problems in Iceland: child outcomes in a nationwide randomized controlled trial. Fam Process. 2015;54(3):498–517.
11. Ren L, Zhang X, Zhou N, Ng ML. Interactive effects between maternal parenting and negative emotionality on social functioning among very young Chinese children. Early Educ Dev. 2017;28(1):21–40. https://doi.org/10.1111/famp.12109.
12. Moffitt TE, Caspi A. Childhood predictors differentiate life-course persistent and adolescence-limited antisocial pathways among males and females. Dev Psychopathol. 2001;13(2):355–75.
13. Burt KB, Obradović J, Long JD, Masten AS. The interplay of social competence and psychopathology over 20 years: testing transactional and cascade models. Child Dev. 2008;79(2):359–74. https://doi.org/10.1111/j.1467-8624.2007.01130.x.
14. Masten AS, Cicchetti D. Resilience in development: progress and transformation. In: Cicchetti D, editor. Developmental psychopathology, vol. 4. 3rd ed. New York: Wiley; 2016. p. 271–333. https://doi.org/10.1002/9781119125556.devpsy406.
15. Domitrovich CE, Cortes RC, Greenberg MT. Improving young children's social and emotional competence: a randomized trial of the preschool "PATHS" curriculum. J Prim Prev. 2007;28(2):67–91. https://doi.org/10.1007/s10935-007-0081-0.
16. McClelland MM, Acock AC, Morrison FJ. The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school. Early Child Res Q. 2006;21(4):471–90. https://doi.org/10.1016/j.ecresq.2006.09.003.
17. OECD. Skills for social progress: the power of social and emotional skills. Paris: OECD; 2015. https://doi.org/10.1787/9789264226159-en.
18. Hupp SD, LeBlanc M, Jewell JD, Warnes E. History and overview. In: Matson JL, editor. Social behavior and skills in children. New York: Springer; 2009. p. 1–21.
19. Sørlie M-A, Amlund-Hagen K, Ogden T. Social competence and antisocial behavior: continuity and distinctiveness across early adolescence. J Res Adolesc. 2008;18:121–44. https://doi.org/10.1080/21683603.2020.1744492.
20. Bornstein MH, Hahn C-S, Haynes OM. Social competence, externalizing, and internalizing behavioral adjustment from early childhood through early adolescence: developmental cascades. Dev Psychopathol. 2010;22(04):717–35. https://doi.org/10.1017/S0954579410000416.
21. Hukkelberg S, Keles S, Ogden T, Hammerstrøm K. The relation between behavioral problems and social competence: a correlational meta-analysis. BMC Psychiatry. 2019;19(1):1–14. https://doi.org/10.1186/s12888-019-2343-9.
22. Bracken BA, Keith LK, Walker KC. Assessment of preschool behavior and social-emotional functioning: a review of thirteen third-party instruments. J Psychoeduc Assess. 1998;16(2):153–69.
23. Merrell KW, Caldarella P. Home and community social behavior scales: user's guide. Baltimore: Paul H. Brookes; 2002.
24. Merrell KW. Using behavior rating scales to assess social skills and antisocial behavior in school settings: development of the school social behavior scales. School Psychol Rev. 1993;22(1):115–33.
25. Hukkelberg S, Ogden T. The Home and Community Social Behavior Scales (HCSBS): dimensionality in social competence and antisocial behaviors. Emot Behav Diffic. 2016;21:377–438. https://doi.org/10.1080/13632752.2016.1201638.
26. Lund J, Merrell KW. Social and antisocial behavior of children with learning and behavioral disorders: construct validity of the home and community social behavior scales. J Psychoeduc Assess. 2001;19(2):112–22. https://doi.org/10.1177/073428290101900201.
27. Merrell KW, Streeter AL, Boelter EW, Caldarella P, Gentry A. Validity of the home and community social behavior scales: comparisons with five behavior rating scales. Psychol Schools. 2001;38(4):313–25. https://doi.org/10.1002/pits.1021.
28. Cordier R, Speyer R, Chen YW, Wilkes-Gillan S, Brown T, Bourke-Taylor H, Doma K, Leicht A. Evaluating the psychometric quality of social skills measures: a systematic review. PLoS ONE. 2015;10(7):e0132299. https://doi.org/10.1371/journal.pone.0132299.
29. Dryburgh NS, Khullar TH, Sandre A, Persram RJ, Bukowski WM, Dirks MA. Evidence base update for measures of social skills and social competence

in clinical samples of youth. J Clin Child Adolesc Psychol. 2020;49(5):573–94. https://doi.org/10.1080/15374416.2020.1790381.

30. Rusch T, Lowry PB, Mair P, Treiblmaier H. Breaking free from the limitations of classical test theory: developing and measuring information systems scales using item response theory. Inf Manag. 2017;54(2):189–203. https://doi.org/10.1016/j.im.2016.06.005.

31. Hukkelberg S, Ogden T. What is social competence? An investigation into the concept among children with antisocial behaviours. Emot Behav Diffic. 2020;25(1):80–93. https://doi.org/10.1080/13632752.2019.1687168.

32. Raimundo R, Carapito E, Pereira AI, Pinto AM, Lima ML, Ribeiro MT. School social behavior scales: an adaptation study of the Portuguese version of the social competence scale from SSBS-2. Span J Psychol. 2012;15(3):1473–84. https://doi.org/10.5209/rev_SJOP.2012.v15.n3.39431.

33. Embretson SE, Reise SP. Item response theory. Hove: Psychology Press; 2013.

34. Ware JE Jr, Bjorner JB, Kosinski M. Practical implications of item response theory and computerized adaptive testing: a brief summary of ongoing studies of widely used headache impact scales. Medical care, II73-II82. 2000. https://www.jstor.org/stable/3768065.

35. Ackerman RA, Donnellan MB, Robins RW. An item response theory analysis of the narcissistic personality inventory. J Personal Assess. 2012;94:141–55. https://doi.org/10.1080/00223891.2011.645934.

36. Van Dam NT, Earleywine M, Borders A. Measuring mindfulness? An item response theory analysis of the Mindful Attention Awareness Scale. Personal Individ Differ. 2010;49(7):805–10. https://doi.org/10.1016/j.paid.2010.07.020.

37. Muraki, E. (1992). A generalized partial credit model: application of an EM algorithm. ETS research report series, no. 1. 1992; p. 1–30. https://doi.org/10.1002/j.2333-8504.1992.tb01436.x.

38. Samejima F. Estimation of latent ability using a response pattern of graded scores. Psychometrika, monograph supplement no. 17. 1969. https://doi.org/10.1002/j.2333-8504.1968.tb00153.x.

39. DeMars C. Item response theory. Oxford: Oxford University Press; 2010.

40. Kjøbli J, Hukkelberg S, Ogden T. A randomized trial of group parent training: reducing child conduct problems in real world settings. Behav Res Ther. 2013;51:113–21. https://doi.org/10.1016/j.brat.2012.11.006.

41. Kjøbli J, Ogden T. A randomized effectiveness trial of brief parent training in primary care settings. Prev Sci. 2012;13:616–26. https://doi.org/10.1007/s11121-012-0289y.

42. Kjøbli J, Ogden T. A randomized effectiveness trial of individual child social skills training: six-month follow-up. Child Adolesc Psychiatry Ment Health. 2014;8(1):31. https://doi.org/10.1186/s13034-014-0031-6.

43. R Core Team. (2019). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/.

44. Chalmers RP. mirt: A multidimensional item response theory package for the R environment. J Stat Softw. 2012;48(6):1–29. https://doi.org/10.18637/jss.v048.i06.

45. Bock RD, Aitkin M. Marginal maximum likelihood estimation of item parameters: application of an EM algorithm. Psychometrika. 1981;46:443–59. https://doi.org/10.1007/BF02293801.

46. Kim SH, Cohen AS, Cho SJ, Eom HJ. Use of information criteria in the study of group differences in trace lines. Appl Psychol Meas. 2019;43(2):95–112. https://doi.org/10.1177/0146621618772292.

47. De Ayala RJ. The theory and practice of item response theory. New York: Guilford Press; 2009.

48. Thissen DE, Wainer HE. Test scoring. Mahwah: Lawrence Erlbaum Associates Publishers; 2001.

49. Maydeu-Olivares A, Joe H. Limited information goodness-of-fit testing in multidimensional contingency tables. Psychometrika. 2006;71:713–32. https://doi.org/10.1007/s11336-005-1295-9.

50. Maydeu-Olivares A, Joe H. Assessing approximate fit in categorical data analysis. Multivar Behav Res. 2014;49(4):305–28. https://doi.org/10.1080/00273171.2014.911075.

51. McDonald RP. Test theory: a unified treatment. Hove: Psychology Press; 2013.

52. Orlando M, Thissen D. Likelihood-based item-fit indices for dichotomous item response theory models. Appl Psychol Meas. 2000;24(1):50–64. https://doi.org/10.1177/01466216000241003.

53. Kang T, Chen TT. Performance of the generalized S-X2 item fit index for polytomous IRT models. J Educ Meas. 2008;45:391–406. https://doi.org/10.1111/j.1745-3984.2008.00071.x.

54. Luo H, Andersson B, Tang JY, Wong GH. Applying item response theory analysis to the montreal cognitive assessment in a low-education older population. Assessment. 2020;27(7):1416–28. https://doi.org/10.1177/1073191118821733.

55. Kim S. A note on the reliability coefficients for item response model-based ability estimates. Psychometrika. 2012;77:153–62. https://doi.org/10.1007/s11336-011-9238-0.

56. Holm S. A simple sequentially rejective multiple test procedure. Scand J Stat. 1979;6(2):65–70.

57. Reise SP, Waller NG. Item response theory and clinical measurement. Annu Rev Clin Psychol. 2009;5(1):27–48. https://doi.org/10.1146/annurev.clinpsy.032408.153553.

58. McFall RM. A review and reformulation of the concept of social skills. Behavioral assessment. 1982.

59. Nangle DW, Grover RL, Holleb LJ, Cassano M, Fales J. Defining competence and identifying target skills. In: Practitioner's guide to empirically based measures of social skills. New York: Springer; 2010. p. 3–19.

60. Dirks MA, Treat TA, Weersing VR. Integrating theoretical, measurement, and intervention models of youth social competence. Clin Psychol Rev. 2007;27(3):327–47. https://doi.org/10.1016/j.cpr.2006.11.002.

61. Rose-Krasnor L. The nature of social competence: a theoretical review. Soc Dev. 1997;6(1):111–35. https://doi.org/10.1111/j.1467-9507,1997.tb00097.x.

62. Youngs D, Canter D, Cooper J. The facets of criminality: a cross-modal and cross-gender validation. Behaviormetrika. 2004;31(2):99–111. https://doi.org/10.2333/bhmk.31.99.

## Publisher's Note