The Lamer Social Competence in Preschool scale (LSCIP): Structural validity in a large

Norwegian community sample

Henrik Daae Zachrisson

The Norwegian Center for Child Behavioral Development & Center for Educational

Measurement at the University of Oslo

Harald Janson

The Norwegian Center for Child Behavioral Development

Kari Lamer

Oslo and Akershus University College of Applied Sciences

Author Note

Henrik Daae Zachrisson, The Norwegian Center for Child Behavioral Development & Center for Educational Measurement at the University of Oslo, email: henrikdz@ulrik.uio.no, Harald Janson, The Norwegian Center for Child Behavioral Development, h.t.janson@nubu.no, Kari Lamer, Oslo and Akershus University College of Applied Sciences, kari@lamer.no. Kari Lamer is Associate Professor Emerita at Oslo and Akershus University College of Applied Sciences.

Corresponding author is Henrik Daae Zachrisson, The Norwegian Center for Child Behavioral Development, PB 7053 Majorstuen, 0306 Oslo, Norway. Email: henrikdz@ulrik.uio.no

ABSTRACT

There is an increasing emphasis on the importance of social competence for children's

development, well-being, and learning. Thus age- and context-appropriate measures are

needed. This study addresses the structural validity of the Lamer Social Competence in

Preschool (LSCIP) scale, developed to accommodate the Nordic model of early childhood

pedagogy. The authors specify the theoretical basis for a bi-factor model of social

competence. This model is tested in a large (n=1157), community based sample of Norwegian

children, including teacher- (at age 2, 3, and 4), mother- (at age 4), and father- (at age 3)

reports. A model with a main factor for social competence, and three bi-factors fit data across

ages and reporters, with factorial, but not strong, invariance over time.

Keywords: Social competence, confirmatory factor analysis, bi-factor model, Lamer Social

Competence in Preschool

2

The Lamer Social Competence in Preschool scale (LSCIP): Structural validity in a large

Norwegian community sample

The purpose of this article is to examine the structural validity of the Lamer Social Competence in Preschool scale (LSCIP; Norwegian name: Lamers skala for sosial kompetanse). This is a relatively widely used measure of social competence in young children in Norway, used both for research purposes and as a pedagogical measure in Early Childhood Education and Care (ECEC) centers. Yet with validation hitherto only presented in a report (Lamer, 2006). Our validation is based on large-scale, longitudinal data with multiple reports, different from the data in which the LSCIP was initially developed and validated.

The development of children's social competence and social skills is an explicit goal of the Nordic model of ECEC (e.g., The Norwegian Ministry of Education, 2017). Moreover, recent work by both educators and economists have emphasized that social competence not only is of vital importance for the child's immediate wellbeing (establishing friendships and dealing with everyday life), it is also a crucial prerequisite for cognitive development and learning, and ultimately for a child's future life chances (e.g., Cunha & Heckman, 2008; Ogden, 2015). This emphasis has had policy implications also beyond the Nordic countries. For example, social competence is one of the core constructs to be assessed in a crossnational study on non-cognitive skills under planning by the Organization for Economic Co-Operation and Development (OECD; Kautz, Heckman, Diris, Weel, & Borghans, 2014). It is also a part of the proposed Early Learning Study by the ECEC Network of OECD (OECD, 2015).

Social competence in children develops over time as children's cognitive and emotional capacities develop and mature, in interaction with the social context (Beauchamp & Anderson, 2010). For example, social communication, perspective taking, and cooperative play are skills with underpinnings in neural development, yet developing in interactions with

other children and adults. Thus, social competence can be assumed to gradually develop throughout the early years, and become a more integrated capacity (Beauchamp & Anderson, 2010). Yet, to our knowledge, there are currently no studies of the normative development of social competences from infancy or toddlerhood, studying the emergence of social competence or it's growth. Some longitudinal studies of the associations between social competence and problem behavior over time, starting around age 4 years, suggest a moderate to strong rank-order stability (.4-.8) throughout school age (Burt & Roisman, 2010; Bornstein, Hahn, & Haynes, 2010).

Despite the emphasis on the importance of social competence, there is little consensus on how to best measure social competence, and whether measures of social competence should be sensitive to the social and cultural context in which children live. A relatively recent systematic review (Humphrey et al., 2011) identified 12 measures of social skills for children and young people (including both measures of skills; i.e., what children do, and competences, i.e., the competences underpinning social skills). Yet, only three of these had been cited in published articles more than 10 times (the Diagnostic Analysis of Nonverbal Accuracy [DANVA; Nowicki & Duke, 1989], Social Competence and Behaviour Evaluation Scale [SCBE; Lafreniere & Dumas, 1996], and Social Skills Rating System [SSRS; Gresham & Elliot, 1990)). All of these were developed in the US, and have occasionally been used in the Nordic countries (e.g., Gamst-Klaussen, Rasmussen, Svartdal, & Strømgren, 2014; Lemola et al., 2011; Sorlie, Hagen, & Ogden, 2008). Validations of these measures in the Nordic countries are uncommon, but the SSIS has been found to meet minimal standards for its psychometric properties in Norway when used in adolescent samples (Gamst-Klaussen et al., 2014; Ogden, 2003). Yet, at a conceptual level, it is questionable whether these measures appropriately measure social competence in preschool children in the way it is fostered and valued in a "Nordic ECEC model". The Nordic ECEC model emphasizes play-based learning, children's free play, active participation by the children in the organization of activities, and

democratic values, as key components. This has been emphasized by for instance Samuelsson & Carlsson (2008), as well as the Norwegian framework plan (i.e., the national "curriculum" for Norwegian ECEC; The Norwegian Ministry of Education, 2017).

As mentioned, the framework plan emphasizes nurturing of social competence as a central theme of the practice in ECEC. Lamer has therefore argued that the conceptualization of social competence in the Nordic context should embrace these values and emphases (e.g., Lamer, 1997a; 2010; 2014).

The development of a Nordic construct of social competence

The LSCIP was designed to measure a Nordic construct of social competence for preschool children. This was initially done to evaluate an intervention aimed at improving social competence in preschool children through the implementation of "You and Me and The Two of Us!" (Norwegian name: "Du og jeg og vi to!"). This program provides a theoretical and methodological framework, including an implementation plan, for increasing children's social competence in ECEC centers (Lamer, 1997a; 1997b; 1997c). The program was designed specifically for the Norwegian context and thus integrated Nordic ECEC values in its design. However, as part of designing the program "You and Me and The Two of Us!", Lamer (1997a) adopted a construct of social competence from the extant research literature, while integrating Nordic ECEC values into a contextually sensitive concept of a social competence. This construct was hence underlying the Lamer Social Competence in Preschool scale (LSCIP). Specifically, Lamer's (1997a) conceptualization was inspired by intervention programs, rating systems and frameworks for preventive mental health: Social Skills Rating System (Gresham & Elliot 1990), ICPS – A Mental Health Program for Kindergarten and First Grade Children (Shure & Spivack 1978), Second Step (Committee for Children, 1989), My Friends and Me (Davis, 1988), and The Competent Child (Strayhorn, 1988). Lamer

integrated the constructs of social competence from these conceptualizations with a critical interpretation of social competence as conceived within the "Nordic model" (Lamer 1997a).

This conceptual analysis resulted in a broad "Nordic construct" of social competence including five sub domains, many of which were inspired by the Social Skills Rating System (Gresham & Elliott, 1990). It is notable that the theoretical model at this point considered these sub domains to be related, and together to constitute social competences. The sub domains were:

- 1) Empathy and role-taking: While empathy is integral to many conceptualizations of social competence for older ages (e.g., Gresham & Elliott, 1990), the construct was here included for preschoolers, and expanded to include role-taking, i.e., the child's capacity to take others' perspectives, values, motives, and needs for communication.
- 2) Pro-social behaviors: Includes children's internalization of fundamental societal values, expressed in behaviors like helping, encouraging, and caring for others.
- 3) Self-control: Includes delay of gratification and impulse control in social interactions, expressed as the ability to wait in turn-taking situations, compromise in conflict situations, in obey common rules and joint decisions, and to plan own behavior in social interactions.
- *Assertiveness:* Addresses children's abilities to take initiative and responsibility, and to be an active part of a social interaction. This sub domain also emphasizes the Nordic values of children's participation and rights to be heard.
- 5) Play, joy, and humor: While this domain is not part of most theories of social competence (e.g., Gresham & Elliott, 1990), play is emphasized in the Nordic ECEC model as a core activity and competence. The domain includes abilities to pretend play, to separate play from reality, to have fun with play, and to enjoy own and other children's mastery.

The development of the Lamer Social Competence In Preschool scale

The LSCIP was developed in a sample of 1426 children, with ages ranging from 1 ½ to 5 years, in 14 ECEC centers participating in at least 2 out of 6 waves of data collection. The sample was part of a trial aimed at testing the intervention "You and Me and The Two of Us!" (Lamer, 2006). For this study, a teacher-report inventory of social competence including 59 items was developed. (Additionally, the first version of the inventory also included brief 6-item scales of externalizing and internalizing behaviors.) The items were theoretically derived from the operationalization of the Nordic construct of social competence as described above, and the inventory therefore included the subscales *empathy* (10 items), *prosocial behavior* (9 items), *self-control* (13 items), *assertiveness* (11 items), *play, joy and humor* (10 items), and *adjustment* (6 items). This latter subscale focused on whether children obey adult instructions and demands in the ECEC centers, like for instance cleaning up after activities, often emphasized in a US context (e.g., Gresham & Elliott, 1990). This latter subscale was not reflected in the program "You and Me and The Two of Us!", as it was already integral to the Norwegian ECEC tradition. Nevertheless, it was still considered to be included in the self-report scale of social competence to increase its comprehensiveness.

The items were subsequently subjected to an exploratory factor analysis, where items were selected based on high convergence with conceptually similar items, combined with low convergence with conceptually dissimilar items, resulting in a subset of 31 items. The selected items covered the five theoretically driven subscales: *empathy and role-taking* (5 items), *prosocial behavior* (5 items), *self control* (6 items), *assertiveness* (6 items), *play, joy, and humor* (5 items), and the additional scale *adjustment* (4 items). For a complete list of these items in Norwegian and English, see Appendix A. The initial psychometric analyses of these six subscales showed high internal consistency within scales (ranging from .89 to .96). Furthermore, exploratory factor analyses showed that the six subscales accounted for 60% to 67% of the total variability across the six waves of data collection. In sum, the development

of the LSCIP was driven by a combination of conceptual and empirical item selection, resulting in a measure with good basic psychometric properties in the initial sample (Lamer, 2006). However, on theoretical grounds, Lamer (2006) combined the following sub scales a) *empathy and role-taking* and *prosocial behavior*, b) *self control* and *adjustment*, and c) *assertion* and *play* in her analyses evaluating the program "You and Me and The Two of Us!". This was done as a parsimonious decision based on in part theoretical considerations; despite being considered as separate aspects of social competence in the literature, these pairs of constructs are thematically overlapping and do reflect similar aspects of both children's behavior and practical pedagogical work in ECEC. In part, the decision was empirically driven; Lamer (2006) found the subscales to be quite highly correlated.

Given the advances in psychometrics, as well as an acknowledgment of the potential for sample-specificity in the psychometric properties of the scale, a more up-to-date psychometric analysis of the scale is warranted. Moreover, given the statistical tools currently more readily available, and the more flexible conceptualization of the construct allowed for by these tools, we aim in the following a reconceptualization of how LSCIP best captures social competence in preschool-aged children.

Structural validity of the Lamer Social Competence In Preschool scale

In a classic chapter on test validation, Cronbach argues that validity is not a property of the test or measure pr. se., but of the interpretation and meaning of the test (Cronbach, 1971). Following this, Messick (1995) proposes an integrated view of construct validity, in which he argues that all aspects of validity and reliability contribute to an adequate interpretation of the measurement score within a context. In this view, construct validity is based on an integration of all evidence supporting a certain interpretation of the test score, and is therefore an evolving process (Messick, 1995). He points to six aspects of construct validity: 1) the content aspect (evidence of content relevance); 2) the substantive aspect

(theoretical rational for observed consistencies in test score); 3) the structural aspect (consistency between expected and observed structure of the construct); 4) the generalizability aspect (generalizability across raters, populations, and settings); 5) the external aspect (convergent and discriminant evidence); and 6) the consequential aspect (valid use of the assessment). These six aspects should be conceived as a heuristic for test validation, rather than a checklist. Expanding on Messick's proposed model for test validation, John and Benet-Martínez (2000) point to the structural aspect by suggesting a design for model testing using Structural Equation Modeling (SEM) for testing Confirmatory Factor Analysis (CFA; Brown, 2006). The main purpose of testing a construct with a measurement model such as CFA is to ensure that our hypothesis about the meaning of the structure in the data (our construct – social competence) reflects the observed structure in the data.

The design of a factor analysis of the LSCIP therefore requires consideration of the meaning we want to infer from the measure. The LSCIP comprises a number of subscales (described above), which measures different components of social competence. Yet, for most applied research purposes, the main interest is in social competence as a broad, and thus unidimensional, construct. This means that a one-unit increase on the measurement scale has the same meaning for each child in the study. A measurement model must therefore be specified in a meaningful way incorporating the multiple dimensions (i.e., the subscales) comprising the broad construct. There are two main ways of specifying such a model (Gignac, 2008); a higher-order (indirect hierarchical) model (see Brown, 2006, for details) and a bi-factorial (direct hierarchical) model (Chen, Hayes, Carver, Laurenceau, & Zhang, 2012). A higher-order model for the LSCIP would include one first-order factor for each of the subscales, and a second-order factor loading on each of the first-order factors. At a conceptual level, this model suggests that the general construct social competence influences each of the subconstructs (e.g., play), which in turn influences the measured scale items (e.g., "is fully

involved in social role play"). Social competence thus causes variation in the items only indirectly, through the first-order factors.

In contrast, the conceptual rational for a bi-factorial, or direct hierarchical, model is different. Social competence is assumed to have a direct influence on each item, while variance at the item level not accounted for by social competence is modeled in additional factors, also loading directly on the item. For example, this model hypothesizes that social competence causes variation in the item "is fully involved in social role play". In addition, other competences also covered by the subscales are also contributing to variability in this item. These could, hypothetically, be imagination and symbolic thinking, i.e., skills that are related to play, but not directly caused by social competence. Thus, a second factor is modelled loading directly on the items related to play, capturing aspects of play which is not caused by social competence. At a conceptual level, we consider this approach be a more meaningful representation of the social competence construct, being competence influencing directly on a broad range of social interactions.

Moreover, for our purpose, there are advantages for model estimation in using a bifactor model over higher-order models (Chen et al., 2012); first it specifies the direct relationship (i.e., strength) between social competence and each item (which is in itself meaningful information, how much variability in "is fully involved in social role play" is in fact accounted for by social competence?); second, it is possible to identify whether a subconstruct (e.g., play) actually exists after accounting for the general construct; third, in applied situations, where the measurement model is part of a structural model, it is easier to separate associations between the main construct (social competence), the sub-constructs, and other variables. In conclusion, we suggest that a bi-factorial model would be both the most meaningful and the most useful way of approaching a structural validation of the LSCIP. In this model, social competence will be loading directly on all the items of the scale, while the six specific subscales will be loading directly on the items conceptually related to those

subscales (and we will have the appropriateness of each sub scale evaluated after accounting for the main construct).

The present study

The purpose of the present study was, in a Norwegian community sample (a separate sample from the one in which LSCIP was initially developed), to test the factor structure of the theoretically derived bi-factorial structure of the LSCIP across multiple time points throughout preschool age, in both preschool teacher and parent reported data. Moreover, we aimed at testing the extent to which LSCIP measures the same construct over time (i.e., measurement invariance), which is a prerequisite for studying change in social competence over time.

Method

Sample and Procedures

Data are from the *Behavior Outlook Norwegian Developmental Study* (BONDS), a prospective, longitudinal study of 1,159 children (559 girls) from 5 municipalities in southeastern Norway. The BONDS is approved by the Regional Committee for Medical and Health Research Ethics and the Norwegian Data Inspectorate, and all parents provided informed written consent. Families were informed about the project in 2006-2008 during their 5-month child health clinic visits in the following five municipalities in the counties of Telemark and Buskerud: Bamble, Porsgrunn, Skien, Tinn, and Drammen. Inclusion criteria were the child being of the appropriate age and at least one parent being able to participate in the study without a translator. Families were informed by the nurse, and provided contact information if agreeing to be contacted. The families of 1,931 eligible children received information. Of these, 1,465 (76%) accepted to be contacted, and 1,159 (79%, or 60% of those originally informed) agreed to participate. The sample fairly well resembles the eligible

families (for complete details, see Nærde, Janson & Ogden, 2014). Two families withdrew their participation and had all data deleted prior to data analysis, reducing the total *n* to 1,157.

Data were collected at multiple time points in the home, and in ECEC centers. Trained assistants conducted interviews with parents in the lab (or home setting, if the parent preferred) at child ages 6 months, 1, 2, 3, and 4 years, which comprised an interview part and a computerized questionnaire section completed by the parent. At 3 and 4 years, the LSCIP was administered to parents as part of the computerized questionnaire procedure. At age 3, fathers were primarily invited to participate, while at age 4, mothers were primarily invited. This rendered responses from 327 mothers and 764 fathers at age 3, and 555 fathers and 1035 mothers at age 4. Moreover, questionnaires including the LSCIP were administered to the ECEC teachers (after parents consented) at child ages 2, 3, and 4 years, requesting that the teacher who knew the target child the best completed the questionnaire. As not all children attended ECEC centers at these ages, and not all centers completed the questionnaires, responses were available from 136, 165, and 159 centers, respectively at ages 2, 3, and 4, covering 65%, 71%, and 60%, the total sample, respectively.

Measures

The Lamer Social Competence in Preschool scale (LSCIP)

The LSCIP is a 31-item adult-report inventory for ECEC teachers or parents including six sub scales: *empathy and role-taking* (5 items), *prosocial behavior* (5 items), *self control* (6 items), *assertiveness* (6 items), *play, joy, and humor* (5 items), and *adjustment* (4 items). Adults are asked to report how often the child shows competencies in the past couple of months. Responses are given on a 5-point Likert-type scale from 1 (*very seldom*) to 5 (*very often*). The inventory in Norwegian and English is presented in Appendix A.

Analyses

We used Mplus, version 7.11 (Muthen & Muthen, 1998-2013) for all inferential analyses. Given the overall normal distribution of the individual item responses, we used the maximum likelihood estimator. The first step of our analyses was to fit a basic model. As a starting point, we used the teacher ratings at age 3 (since this was the middle time points of our assessments), and subsequently fitted the preferred model from age 3 for each of the assessed time points (ages 2, 3, and 4) and reporters (preschool teachers, mothers, and fathers). This stage of the modeling process is also known as configural invariance (i.e., is the factor structure identical across time and reporters, indicative of whether the same construct is measured). Since the LSCIP is a new measure, and has not previously been subjected to a structural validation, this step involved an iterative process where we tested the adequacy of items and sub scales for an overall assessment of a social competence construct. Our next step involved constraining factor loadings across time points to test for longitudinal metric invariance (i.e., are the factor loadings similar across time, indicative of whether the respondents attribute the same meaning to construct over time), and constraining intercepts to test scalar invariance, to determine whether means can be compared over time.

Model fit was evaluated according to conventional standards for fit indexes, with Root Mean Square Error of Approximation (RMSEA) below .06, Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) above .95 indicating very good model fit, and values below .08 and above .9, respectively, indicating reasonable fit (Hu & Bentler, 1999).

We also examined model-based reliability and dimensionality of the final model. In doing this, we follow recommendations by Rodrigues, Reise, and Haviland (2016).

Results

Initial model-fitting

Our first analytic step was to determine a baseline bi-factor CFA-model in teacher-reported data at age 3. We chose teacher reports at 3 years because the LSCIP was designed to measure social competence in a preschool setting, and hence teachers were the intended reporters of the scale. Moreover, as 3 years is the middle age for our teacher reports, thus serving a meaningful baseline for a measurement model to be expanded to 2 and 4 year olds. To recapture, this model hypothesizes that there is one main social competence factor loading on all the 31 indicators, and that there are six additional factors, reflecting each of the original subscales.

For comparative purposes, we first fitted a one-factor model, with all 31 items loading on one social competence factor. This model fitted poorly (Chisq [434]=5353, p<.000, RMSEA=.117, CFI/TLI=.632/.606). Moreover, modification indices suggested a vast number of residual correlations. While 28 factor loadings ranged from .411 to .774, three items had low factor loadings (.129, .202, .333), for the items "Reacts critically to rules that are perceived as unfair", "Speaks out clearly when s/he conceives something as unfair", and "Can resist group pressure", respectively. These items were initially in the assertion subscale, but were dropped from further analyses.

Second, we fitted a full bi-factor model, with one social competence factor loading on the remaining 28 items, and in addition a bi-factor for each of the theoretically derived subscales loading on their respective items (see Table 1). As could be expected, this over-identified model failed to converge. We assumed the subscale prosocial behavior to be most uniquely predicted by social competence. Because of this, we fitted a second bi-factor model with five bi-factors, while removing the prosocial behavior bi-factor, assuming social competence to cause all meaningful variation in these items. Model fit improved (Chisq [327]=1558, *p*<.000, RMSEA=.067, CFI/TLI=.903/.888), while still indicating mediocre model fit according to the CFI and TLI indices. We thus refitted a model based on the theoretical and empirical work by Lamer (2006), collapsing the subscales *self-control* and

adjustment into one bi-factor, and assertion and play into another, while maintaining the subscale *empathy and role-taking* as a separate bi-factor. Informed by results from the previous models, where items loading on prosocial behavior had the highest factor loadings, we considered this factor to be most uniquely aligned with the over-all social competence factor. For details, see Table 1. Finally, informed by modification indices in previous models, we allowed for residual correlations due to shared specific content between the items "Helps other children without being asked for it" and "Helps you without being asked for it", and also between the items "Completes tasks she/he is asked to do" and "Completes tasks she/he is given within designated time". This final model fitted data adequately in teacher reported data at age 3 (Chisq [325]=1378, p<.000, RMSEA=.063, CFI/TLI=.917/.903). Factor loadings for this model are displayed in Table 2. As the model fitting process was in part conceptually, in part empirically guided, we consider it a meaningful and adequately fitted measurement model. Additionally, as a robustness check, we re-estimated this model including the three items initially removed from the model. This 31 item model had worse model fit than the reduced 28 item model (Chisq [405]=1921, p < .000, RMSEA=.067, CFI/TLI=.887/.870), and we therefore maintained the reduced number of items (28) for further analyses.

Robustness of the bi-factor model across ages and reporters

Our next goal was to test the final bi-factor model we described above on teacher reports at ages 2 and 4, and on father reports on age 3 as well as mother reports on age 4. Fit indexes for these models can be seen in Table 2. For teacher reports, the models fitted adequately at ages 2 and 4, except for the TLI at age two which was slightly below the recommended cut-off at .9. The same was the case for mother report at age 4. For father report at age 3, however, both CFI and TLI values were below the recommended cut-off.

Does LSCIP measure the same construct over time?

Our final aim was to test whether the bi-factor model of LSCIP measured the same social competence construct over time, i.e., to test for measurement invariance, in teacher

reports for which we had repeated measures from ages 2, 3, and 4 years. In doing this, we followed the steps outlined in the Analysis section above. We started with examining the baseline model (i.e., configural invariance), including teacher reports at ages 2 and 3. The purpose was to establish whether the same factor model could be applied across these two ages.

Fit indexes for the invariance analyses are displayed in Table 4. For ages 2 and 3, the unconstrained model (i.e., with all factor loadings allowed to vary freely across ages, and standardized variance in the latent factors) showed adequate fit. This supports the interpretation from our previous analyses, that the bi-factor model fits well across these two age groups. The next step was to constrain the factor loadings to be equal across time (i.e., so the loading on the first item at age 2 was equal to the loading on the first item at age 3). This lead to negligible reductions in fit indexes, indicating that the measure has weak factorial (a.k.a. metric) invariance across these ages. This means that the covariance across these two ages is due to the global social competence factors, that is, that the association between social competence at ages 2 and 3 can be estimated (e.g., as part of cross-lagged models). Weak factorial invariance does, however, not allow comparisons of latent means over time, that is, conclusions about whether social competence increases or decreases over time, as this requires comparable latent means. We tested this in our next step, constraining the intercepts of the factor loadings to be equal over time, while freely estimating the latent mean at age 3. As can be seen in Table 4, this model showed considerable reduction in CFI and TLI, suggesting that latent means cannot be compared across these ages, nor should models requiring estimates of means (e.g., latent growth curves) be applied using this measurement model. In these analyses, we constrained the intercepts for social competence factor only, while leaving the loadings for the bi-factors to be freely estimated.

We repeated these analyses across ages 3 and 4 years, and across 2, 3, and 4 years (see Table 4 for fit indexes). The results showed similar pattern, with evidence for weak factorial invariance, but not for strong (a.k.a. scalar) invariance over time. Suspecting that some items could be more strongly subject to change due to development, we inspected the residuals from the freely estimated models (see Appendix B). While most intercepts increased across time, three items (22, 27, 29) had considerable increase, above .5 of an unstandardized scale score. We reran the invariance models without constraining these three intercepts over time, without improvements of model fit. Thus, the lack of strong measurement invariance seemed to apply regardless of these specific items.

Model-based reliability and dimensionality

We then performed additional tests of the final bi-factor model, as recommended by Rodriguez et al. (2016). We did this for the teacher-rating at age 3, as we considered this to be our primary model. The model-based reliability measures whether the LSCIP scores represent the constructs of interest (see also Hammer & Toland, 2016).

The omega for the total score, i.e., the proportion of total variance explained by the entire model, was .95. The hierarchical omega for the social competence factor, that is, the proportion of total score variance that can be attributed to this factor accounting for the bifactors, was .82. The hierarchical omega for the three bi-factors were .04, .08, and .01, for self-control and adjustment, assertion and play, and empathy and role-taking, respectively. This indicates that the total score of the LSCIP predominantly reflects one single factor, while the majority of the subscale score variances are due to the general social competence factor, and thus do not reliably measure the subdomain constructs. We then estimated the percentage of reliable variance in the social competence factor independent of the specific factors, which was 86%. For the subscales, the proportion of reliable variance for the specific factors independent of the social competence factor was 37%, 53%, and 29%, respectively. Again, the subscales seem unreliable as unique measures of the bi-factors. In sum, the reliability estimates suggest that while the bi-factor model can be used as a measure for social

competence, the bi-factors should not be used as measures of self-control and adjustment, assertion and play, and empathy and role-taking, respectively.

In order to examine the unidimensionality of the LSCIP, we calculated the explained common variance (ECV), which is the proportion of common variance across items that can be explained by the social competence factor. EVC for the social competence factor was .63. This suggests that the LSCIP is multidimensional, and that the use of the measure should account for the bi-factors. Yet, the percent of uncontaminated correlations (calculated using Hammer's [2016] online calculator) was .67, which, when the hierarchical omega is above .70, does not disqualify the interpretation of the measure as unidimensional (Hammer & Toland, 2016). In sum, the social competence factor may be considered a unidimensional measure when accounting for the additional bi-factors. Finally, we saved out a factor score of the main social competence factor (teacher rated at 36 months), and correlated it with a mean score of the 31 item version. The two scales were almost perfectly correlated (r = .999), and variance was practically identical.

Discussion

The purpose of the present study was to test the structural validity of the Lamer Social Competence in Preschool (LSCIP) scale in a large Norwegian community sample. The LSCIP was designed to encompass a construct of social competence valued by the Nordic ECEC model, with its play-based and democratic approach to child development. In our conceptual analysis, we argued that a bi-factor model was appropriate for providing a unidimensional construct based on a measure consisting of six subscales. Through a number of steps, our theoretical bi-factor model was modified to fit adequately to teacher-ratings at age 3, and subsequent tests of the model using teacher ratings at age 2 and 4 (within the same sample) yielded fairly similar model fit. The model also fit mothers' ratings at age 4, but less well fathers' ratings at age 3. Moreover, we found evidence for factorial invariance for teacher

ratings over time, but not for scalar invariance over time. The main social competence factor was reliable, while the bi-factors were not. Thus, the bi-factors should not be used as dependent variables in subsequent analyses. Our additional analyses of dimensionality suggest that when accounting for the bi-factor structure, the LSCIP is a unidimensional measure of social competence. In sum, our analyses suggest that the LSCIP is a comprehensive measure of social competence in the preschool years. The bi-factor model of the LSCIP (after deleting three items) provides a unidimensional measure of the construct when rated by teachers throughout early childhood, and by mothers at age 4, but caution should be applied when applied to fathers at age 3. Moreover, the measure could be used with confidence in longitudinal structural models of social competence, but analyses of changes in mean values of social competence through the preschool years were not supported by our models.

In our theoretical analysis of the social competence construct as conceptualized by the LSCIP, we argued for a bi-factor model over a higher-order model. To recapture, our thinking was that social competence is conceived as a broad and comprehensive, yet uni-dimensional, construct. In other words, social competence is by many conceived to vary along one dimension, some children have less of this competence than other children, yet social competence is expressed in different types of situations, like play, empathy, etc. The child's behavior in these situations may, however, also be influences by other competences, for example, play may also be influenced by a child's imagination and symbolic thinking. We therefore argued that modeling this type of dual influence on behaviors rated at the item-level would be the best way of capturing true differences in social competence, while (in the model) acknowledging the complexities involved in a theoretical understanding of social behavior. Our findings were partially in support of this thinking. Through the model-fitting process, we made adjustments to the bi-factors in accordance with the theoretical model proposed by Lamer (2006). We did this by reducing the number of bi-factors to three, in part by collapsing the play and assertion subscales, and the self-control and adjustment subscales.

We also kept the empathy/ role-taking-subscale as a third bi-factor. Our final bi-factor model had rather strong loadings on the main factor (median factor loading = .569, range .309-.818), with weaker loadings for the bi-factors. This final model suggests that while variability in all items (except for the three which we deleted initially) are caused by social competence, additional variability in the items belonging to these subscales are accounted for by other constructs.

The implications of these findings for the practical use of the LSCIP for researchers and practitioners are both encouraging and cautionary. While the bi-factor model is a sound model for use in latent variable analyses, the almost perfect correlation between the factor score and the mean scale score enables the more convenient use of the mean score rather than a latent model in measurement of one group of children at one time point. Moreover, we found partial support for the notion that the LSCIP, although designed to be used in ECEC settings, may also be used as a parent-rated scale, at least when mothers are asked to report. Mothers and ECEC teachers rate indicators of social competence in largely the same way, evident by the fact that the factor model fitted ratings by both these groups. We have no clear idea about why the structure of fathers' ratings of social competence items does not correspond exactly with that of mothers' and teachers' ratings, reflected in poorer model fit. A number of explanations may be hypothesized for the finding of poorer correspondence of the model for fathers. Fathers of three-year-olds may possibly have less experience of their child in situations where social competence is elicited. There may also be reporter-gender-bias for children's social competence (most ECEC teachers are women and might, assuming such a hypothesis, be expected to be more similar to mothers than to fathers in their responses).

The final, and very important, practical implication of our finding is that the LSCIP cannot not be used to measure whether children display increased levels of social competence in ECEC across ages 2 through 4. Our invariance analyses suggest that the theoretical structure of social competence, as we tested it, is identical over time, but that latent measure

means—and hence manifest item means—seem to be different across ages. One plausible reason for such a finding is that adults implicitly apply some age-referencing when they respond to items, possibly both with respect to frequency (i.e., judging whether a child does a behavior seldom or often in relation to what is seen as appropriate for age) and behavior content (i.e., thinking of age-appropriate instances of behaviors such as initiating contact, participating in play, or supporting other children, when judging frequency). Thus, it is meaningful to analyze longitudinal data of associations between social competence and other constructs over time. Yet, our analyses do not support the comparison of means over time. This includes analyses of changes in social competence, e.g., growth curve analyses. However, the scale score may be used to quantify between-group differences in increase in social competence over time, for example between an intervention group and a control group, given that between group invariance pre and post intervention is established. It is, of course, yet to be tested whether this finding replicates in different samples and populations. The preliminary conclusion is therefore that if assumptions of age-invariance are made in future studies, they must be formally tested before being applied.

This study has some notable strengths. We contribute with a rigorous psychometric test of a measure of social competence developed with a particular pedagogical and socio-cultural context in mind. We have a large sample, followed longitudinally with high participation rate and low attrition, from a demographically rather diverse population, reflecting normal children in Norwegian ECEC centers and families. We therefore argue that the ecological validity of this study is high. However, some limitations apply. Our sample sizes for mothers and fathers differed, and selection bias in which fathers actually participated may have influenced the results. Given the design of the BONDS study, with mothers and fathers being targeted in different waves of the data collection, we were prevented from conducting longitudinal analyses of these respondents. Finally, given that our sample was recruited from

the normal population, our results may not directly apply to selected groups (e.g., special needs populations).

The generalizability of our findings to other samples remains an empirical question.

Rather than being conclusive, our findings should be considered guiding for future use of the LSCIP in being sensitive to measurement issues, while at the same time being optimistic that the LSCIP is a context sensitive scale of social competence. Future research should not only strive to replicate or modify our measurement model, but also to expand the evidence for the validity of the LSCIP to include theoretically meaningful relations to other related constructs. For instance, longitudinal relations to later social competence in school-age, associations with behavior problems and other domains of child functioning and well-being, and ECEC and family predictors of social competence in the preschool years are valuable applications of the LSCIP in future research. The fact that we found mother reports to fit the same measurement model as teacher reports should encourage the use of the LSCIP also in this population.

In conclusion, this study provides evidence that the LSCIP can be used as a measure of social competence in Norwegian preschool children. However, we strongly encourage future use of the scale to be sensitive to modelling issues, especially when used in longitudinal analyses.

References

- Beauchamp, M. H., & Anderson, V. (2010). SOCIAL: An integrative framework for the development of social skills. *Psychological Bulletin*, *136*, 39-64. doi:10.1037/a0017768
- Bornstein, M. H., Hahn, C., & Haynes, O. M. (2010). Social competence, externalizing, and internalizing behavioral adjustment from early childhood through early adolescence:

 Developmental cascades. *Development and Psychopathology*, 22(4), 717-35.

 doi:http://dx.doi.org/10.1017/S0954579410000416
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Publications.
- Burt, K. B., & Roisman, G. I. (2010). Competence and psychopathology: Cascade effects in the NICHD study of early child care and youth development. *Development and Psychopathology*, 22(3), 557-67. doi:http://dx.doi.org/10.1017/S0954579410000271
- Chen, F. F., Hayes, A., Carver, C. S., Laurenceau, J. P., & Zhang, Z. (2012). Modeling general and specific variance in multifaceted constructs: A comparison of the bifactor model to other approaches. *Journal of Personality*, 80, 219-251. doi: 10.1111/j.1467-6494.2011.00739.x
- Committee for Children 1989. Second Step. A Violence-prevention Curriculum. Seattle, Washington.
- Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), *Educational Measurement* (2nd ed.). Washington, D. C.: American Council on Education. (Reprinted from: Not in File).
- Cunha, F., & Heckman, J. J. (2008). Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Journal of Human Resources*, 43, 738-782. doi:10.1353/jhr.2008.0019
- Davis, D.A. 1988. My Friends and Me. American Guidance Service, Circle Pines, Minn.

- Gamst-Klaussen, T., Rasmussen, L.-M. P., Svartdal, F., & Strømgren, B. (2014).
 Comparability of the Social Skills Improvement System to the Social Skills Rating
 System: A Norwegian Study. Scandinavian Journal of Educational Research (ahead-of-print), 1-12. doi: 10.1080/00313831.2014.971864
- Gignac, G. E. (2008). Higher-order models versus direct hierarchical models: g as superordinate or breadth factor? *Psychology Science Quarterly*, 50, 21-43.
- Gresham, F. M., & Elliott, S. N. (1990). *Social skills rating system (SSRS)*: American Guidance Service.
- Hammer, J. H., & Toland, M. D. (2016, November). *Bifactor analysis in Mplus*. [Video file]. Retrieved from http://sites.education.uky.edu/apslab/upcoming-events/
- Hammer, J. H. (2016, October). Percent of Uncontaminated Correlations (PUC) Calculator: A Microsoft Excel-based tool to calculate the PUC statistic. Available from http://DrJosephHammer.com/.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexies in covariance structure analysis: Convensional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55. doi:10.1080/10705519909540118
- Humphrey, N., Kalambouka, A., Wigelsworth, M., Lendrum, A., Deighton, J., & Wolpert, M.
 (2011). Measures of Social and Emotional Skills for Children and Young People A
 Systematic Review. *Educational and Psychological Measurement*, 71, 617-637. doi: 10.1177/0013164410382896
- John, O. P., & Benet-Martinez, V. (2000). Measurment: Reliability, construct validation, and scale construction. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods* in social and personality psychology (pp. 339-369). Cambridge: Cambridge University Press.
- Kautz, T., Heckman, J.J., Diris, R., Weel, B.t., & Borghans, L. (2014). Fostering and

 Measuring Skills: Improving Cognitive and Non-cognitive Skills to Promote Lifetime

- Success. OECD Education Working Papers, No. 110, OECD Publishing, Paris. DOI: 10.1787/5jxsr7vr78f7-en
- LaFreniere, P. J., & Dumas, J. E. (1996). Social competence and behavior evaluation in children ages 3 to 6 years: The short form (SCBE-30). *Psychological Assessment*, 8, 369-377. doi: 10.1037/1040-3590.8.4.369
- Lamer, K. (1997a). Du og jeg og vi to! Om å fremme barns sosiale kompetanse. Teoriboka.

 [You and Me and The Two of Us! About developing childrens social competence. The theory book]. Oslo: Gyldendal Akademisk.
- Lamer, K. (1997b). Du og jeg og vi to! Et rammeprogram for sosial kompetanseutvikling.

 Håndboka [You and Me and The Two of Us! A framework program for development of social competence. The handbook]. Oslo: Gyldendal Akademisk.
- Lamer, K. (1997c). Du og jeg og Truls og Trine! Barneboka [You and me and Truls and Trine! The children's book]. Oslo: Gyldendal Akademisk.
- Lamer, K. (2006). Evaluering og videreutvikling av "Du og jeg og vi to!" [Evaluation and further development of "Du og jeg og vi to"]. In K. Lamer & S. Hauge (Eds.), Fra rammeprogram til handling [From framework plan to action] HiO-rapport nr. 28.

 Oslo, Norway: Oslo and Akershus University College of Applied Sciences. pp. 15-57.
- Lamer, K. (2010). Utvikling av sosial kompetanse og inkludering i barnehagen som lærende organisasjon [Development of social competence and inclusion in the child care center as a learning organization], i Ø. Kvello (Ed.): *Barnas barnehage 2. Barn i utvikling*, Oslo: Gyldendal akademisk. pp. 229 255.
- Lamer, K. (2014). Dette vet vi om barnehagen: Sosial kompetanse [This is what we know about center care: Social competence]. Oslo, Norge: Gyldendal Akademisk.
- Lemola, S., Räikkönen, K., Scheier, M. F., Matthews, K. A., Pesonen, A. k., Heinonen, K., . . . Kajantie, E. (2011). Sleep quantity, quality and optimism in children. *Journal of Sleep Research*, 20, 12-20. doi: 10.1111/j.1365-2869.2010.00856.x

- Messick, S. (1995). Validity of psychological assessment. Validation of inferences from persons' responses and performances. A scientific inquiry into score meaning.

 *American Psychologist, 50, 741-749. doi.org/10.1037/0003-066X.50.9.741
- Ministry of Education. (2017). Rammeplan for barnehagen: Forskrift om rammeplan for barnehagens innhold og oppgave. [Framework plan for center-based child care: Regulation concerning a framework plan for content and duties of center-based child care.] Oslo, Norway: Kunnskapsdepartementet [Ministry of Education and Research], 2017.
- Muthen, L. K., & Muthen, B. O. (1998-2013). Mplus (version) 7.11 [Computer Software]. .

 Los Angeles, CA: Muthén & Muthén.
- Nowicki, S., Jr., & Duke, M. P. (1989). A measure of nonverbal social processing ability in children between the ages of 6 and 10. Paper presented as part of a symposium at the American Psychological Society, Alexandria, VA.
- Organisation for Economic Co-operation and Development (OECD) (2015). Call for tenders:

 International Early Learning Study. Available at:

 http://www.oecd.org/callsfortenders/CfT%20100001420%20International%20Early%
 20Learning%20Study.pdf
- Ogden, T. (2003). The validity of teacher ratings of adolescents' social skills. *Scandinavian Journal of Educational Research*, 47, 63-76.
- Ogden, T. (2015). Sosial kompetanse og problematferd blandt barn og unge [Social competence and problem behavior among children and adolescents]. Oslo, Norway: Gyldendal Akademisk.
- Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models:

 Calculating and interpreting statistical indices. *Psychological Methods*, *21*, 137-150.

 doi:10.1037/met0000045

- Samuelsson, I. P., & Carlsson, M. A. (2008). The Playing Learning Child: Towards a pedagogy of early childhood. *Scandinavian Journal of Educational Research*, *52*, 623-641. doi:10.1080/00313830802497265
- Shure, M.B. og G. Spivack 1978. *ICPS A Mental Health Program for Kindergarten and First Grade Children*. Rev. utg. Hahneman University, Department of mental health science, Philadelphia.
- Sorlie, M. A., Hagen, K. A., & Ogden, T. (2008). Social competence and antisocial behavior:

 Continuity and distinctiveness across early adolescence. *Journal of Research on Adolescence*, 18, 121-144. doi: 10.1111/j.1532-7795.2008.00553.x
- Strayhorn, J.M. 1988. *The Competent Child. An approach to psychotherapy and preventive mental health.* New York: The Guilford Press.

Item wordings and factor loadings (standard errors) for the final base-line bi-factor model of LSCIP, teacher reported at age 3. N

Item w	Item wordings and factor loadings (standard errors) for the final base-line bi-factor model of LSCIP, teacher reported at age 3. N					
Item	Wording	Sub	Social	Play/	S	
#		Scale	Competence	Assertion		
17	Supports and encourages the other children	PS	0.818 (0.014)			
29	Says something nice ^a	PS	0.731 (0.018)			
5	Helps the other children without being asked ^b	PS	0.717 (0.019)			
23	Helps other children in conflict situations	PS	0.702 (0.020)			
11	Helps you without being asked ^b	PS	0.658 (0.022)			
16	Initiates play	PL	0.537 (0.027)	0.668(0.024)		
10	On her/his own initiative joins in other children's play ^a	PL	0.525 (0.027)	0.660(0.024)		
4	Wants to participate in play or other group activities	PL	0.526(0.027)	0.530(0.027)		
22	Involves her/himself completely in social role play	PL	0.590 (0.025)	0.414(0.414)		
28	Makes friends easily	PL	0.622 (0.024)	0.408(0.029)		
1	Initiates contact (in an OK manner)	AS	0.611 (0.024)	0.334(0.030)		
13	Meets new people with openness, makes eye contact	AS	0.439 (0.030)	0.250(0.035)		
7	Speaks when others are present (in an OK manner)	AS	0.645 (0.022)	0.219(0.031)		
14	Can control anger in conflicts with the other children	SC	0.315 (0.033)		0.7	
31	Can control her/his anger in conflicts with adults	SC	0.309 (0.034)		0.6	
2	Accepts that her/his wishes will not always be fulfilled	SC	0.369 (0.032)		0.6	
20	Adjusts ^a	SC	0.391 (0.032)		0.6	
8	Waits for her/his turn in games and other activities	SC	0.538 (0.027)		0.4	
26	Compromises in conflict situations ^a	SC	0.502 (0.028)		0.4	
6	Does as she/he is asked	AD	0.527 (0.027)		0.4	
12	Completes tasks she/he is assigned ^c	AD	0.569 (0.026)		0.2	
18	Completes tasks she/he is given within designated time ^c	AD	0.601 (0.024)		0.2	
24	Cleans up after her/himselfa	AD	0.418 (0.031)		0.2	
15	Shows that she/he sees that others are angry	EM	0.598 (0.024)			
9	Shows that she/he sees that others are sad	EM	0.635 (0.023)			
21	Shows that she/he sees that others are afraid	EM	0.550(0.027)			
3	Shows that she/he sees that other are happy	EM	0.639 (0.023)			
27	Recognizes, and can put words, to others' feelings	EM	0.633 (0.023)			

Note. ^aFor full item wording, see Appendix A. ^bCorrelated errors, 0.293. ^cCorrelated errors, 0.480. EM= Empathy and role-taking; SC=Self-control; AS=Assertiveness; PL=Play, joy, and humor; AD=Adjustment

Table 2. Fit indexes for the bi-factor model of LSCIP

Reporter and age	N	Chisqb	RMSEA	CFI	TLI	Factor loadings
						(Range ^a)
Teacher report, age 2	751	1292	0.063	0.906	0.891	0.229-0.792
Teacher report, age 3	828	1378	0.063	0.917	0.903	0.309-0.818
Teacher report, age 4	700	1337	0.067	0.917	0.903	0.423-0.837
Father report, age 3	764	989	0.052	0.897	0.880	0.275-0.769
Mother report, age 4	1035	1337	0.055	0.907	0.892	0.314-0.814

Notes: ^afor the global social competence factor. ^bAll chi-square tests were significant at *p*<.001

Table 3. Fit indexes for tests of measurement invariance

Reporter and ages	Constraints	Chisq ^a (DF)	RMSEA	CFI	TLI
Teacher report, age 2-3	None	3552 (1403)	0.041	0.909	0.900
	Factor loadings	3626 (1431)	0.041	0.907	0.900
	+ Intercepts	4388 (1460)	0.048	0.868	0.861
Teacher report, age 3-4	None	3619 (1403)	0.041	0.914	0.905
	Factor loadings	3626 (1431)	0.042	0.911	0.905
	+ Intercepts	4186 (1460)	0.048	0.881	0.875
Teacher report, age 2-4	None	6711 (3231)	0.033	0.905	0.898
1	Factor loadings	6896 (3288)	0.033	0.902	0.896
	+ Intercepts	8487 (3343)	0.040	0.860	0.854

Notes: ^aAll chi-square tests were significant at p<.001.

APPENDIX A

Lamer Social Competence in Preschool Scale Lamers sosial kompetanseskala

	English translation	Norwegian original
	Social competence	Sosial kompetanse
	We want to learn something about how often	Vi ønsker å få vite noe om hvor ofte barnet viser
	the child shows certain social competences.	bestemte sosiale ferdigheter. Les hver av utsagnene
	Read each of the statements below and think	nedenfor og tenk gjennom atferden til dette barnet i
	through the behavior of the child during the last	løpet av de siste par månedene. Kryss av for hvor
	couple of months. Check how often you think	ofte du mener at barnet viser den aktuelle
	that the child shows the specific competence.	kompetansen.
	that the child shows the specific competence.	Kompetansen.
	Response Categories	Svarkategorier
1	Very seldom	Svært sjelden
2	Seldom	Sjelden
3	Sometimes	Av og til
4	Often	Ofte
5	Very often	Svært ofte
	T.	* 11
1	Items Initiates contact (in an OK manner)	Ledd Tar initiativ til kontakt (på en OK måte)
2	Accepts that her/his wishes will not always be	Aksepterer at egne ønsker ikke alltid blir oppfylt
2	fulfilled	Aksepterer at egne ønsker ikke antid om opprytt
3	Shows that she/he sees that others are happy	Viser at hun/han ser at andre er glade
4	Wants to participate in play or other group	Har lyst til å delta i lek eller andre gruppeaktiviteter
	activities	,
5ª	Helps the other children without being asked	Hjelper de andre barna uten å bli bedt om det
6	Does as she/he is asked.	Gjør det du ber hun/han om
7	Speaks when several others are present (in an	Tar ordet når flere er samlet (på en OK måte)
	OK manner)	
8	Waits for her/his turn in games and other	Venter på tur i spill og andre aktiviteter
	activities	
9	Shows that she/he sees that others are sad	Viser at hun/han ser at andre er lei seg
10	On her/his own initiative joins in other	Tar selv initiativet til å bli med på andre barns lek
	children's play or activities	eller aktiviteter
11	Helps you without being asked	Hjelper deg uten å bli bedt om det
12	Completes tasks she/he is assigned	Fullfører oppgaver som det blir satt til
13	Meets new people with openness, makes eye	Møter nye mennesker med åpenhet, tar øyekontakt
	contact	
14ª	Can control anger in conflicts with the other	Kan styre sinnet sitt i konflikter med de andre barna
1.5	children	Visas et lass /lass ess et au dus en sinta
15	Shows that she/he sees that others are angry	Viser at hun/han ser at andre er sinte
16	Initiates play	Tar initiativ til lek
17a	Supports and encourages the other children	Støtter og oppmuntrer de andre barna
18	Completes tasks she/he is given within the	Gjør de oppgavene hun/han får innen avsatt tid
19 ^b	designated time Reacts critically to rules that are perceived as	Reagerer kritisk på regler som oppleves som
19	unfair	urettferdige
20	Adjusts (gives in, adapts her/himself, admits	Jenker seg (gir seg, innordner seg, innrømmer egne
20	own errors, forgives others)	feil, tilgir andre)
21	Shows that she/he sees that others are afraid	Viser at hun/han ser at andre er redde
22	Involves her/himself completely in social role	Involverer seg fullt og helt i sosial rollelek
22	play	involverer seg runt og heit i sosiar forterek
23	Helps other children in conflict situations	Hjelper andre barn i konfliktsituasjoner
24	Cleans up after her/himself when play/activities	Rydder opp etter seg når leken/aktiviteten avsluttes
∠ ⊤	are terminated	rejuder opp etter seg har rekentaktiviteten avsluttes
25 ^b	Can resist group pressure	Kan stå imot gruppepress
26	Compromises in conflict situations (e.g., by	Inngår kompromisser i konfliktsituasjoner, f.eks. ved
20	changing own opinions or adjusting own	å endre på egne meninger eller fire på egne ønsker
	wishes)	a chare pa egue meninger ener me pa egue pusker

27	Recognizes, and can express in words, others'	Gjenkjenner, og kan sette ord på, andres følelser		
28	feelings Makes friends easily	Får lett venner		
_	•	1 00 1000 / 0111101		
29	Says something nice, gives compliments, to the other children	Sier noe hyggelig, gir komplimenter, til de andre barna		
30 ^b	Speaks out clearly when she/she conceives something as unfair	Sier tydelig ifra når hun/han synes noe er urettferdig		
31	Can control her/his anger in conflicts with adults	Kan styre sinnet sitt i konflikter med voksne		
	^a Adjustments to the	parent rated version		
5	Helps other children without being asked	Hjelper andre barn uten å bli bedt om det		
14	Can control anger in conflicts with other children	Kan styre sinnet sitt i konflikter med andre barn		
17	Supports and encourages other children	Støtter og oppmuntrer andre barn		
29	Says something nice, gives compliments, to other children	Sier noe hyggelig, gir komplimenter, til andre barn		

^bThese items were removed in the final models because of low factor loadings.

Notes about the translation:

The translation into English was done by the first author (INSERT NAME AFTER REVIEW). A clinical psychologist who is native speaker in both English and Norwegian back-translated. Edits of the English translation was done by the first and second authors (INSERT NAME AFTER REVIEW) based on the back – translation. Items 7 and 10 were rephrased after back-translation, otherwise the initial translations were kept with minimal edits. More information about the translation is available from the authors.

APPENDIX B

Intercepts for the teacher rated LSCPI items from the unconstrained models ages 2, 3, and 4 years.

Item#	Age 2		Age 3		Age 4	
	Est	SE	Est	SE	Est	SE
2	3.366	0.029	3.466	0.026	3.669	0.030
2 3	3.992	0.029	3.983	0.026	4.046	0.029
4	4.269	0.029	4.367	0.026	4.409	0.027
5	3.067	0.035	3.257	0.033	3.453	0.036
6	3.705	0.027	3.724	0.026	3.912	0.029
7	3.196	0.034	3.527	0.030	3.892	0.031
8	4.102	0.029	4.223	0.026	4.310	0.026
9	3.095	0.042	3.445	0.035	3.679	0.036
10*	3.766	0.038	3.670	0.035	3.787	0.037
11	3.702	0.032	3.831	0.026	3.866	0.031
12	3.788	0.034	4.026	0.028	4.192	0.029
13	2.872	0.036	3.090	0.033	3.216	0.037
14	3.284	0.031	3.498	0.029	3.755	0.031
15	3.245	0.037	3.469	0.031	3.690	0.036
16	3.437	0.032	3.650	0.026	3.774	0.030
17	3.947	0.034	4.185	0.029	4.322	0.029
18	2.923	0.036	3.167	0.031	3.314	0.034
19	3.067	0.034	3.335	0.029	3.600	0.031
20	3.067	0.032	3.134	0.029	3.295	0.031
21	2.948	0.034	3.209	0.030	3.427	0.033
22ª	2.979	0.043	3.799	0.036	4.120	0.036
23	2.344	0.034	2.752	0.031	3.026	0.034
24	2.891	0.036	2.997	0.031	3.276	0.036
26	2.715	0.035	2.895	0.027	3.116	0.031
27^{a}	3.002	0.043	3.562	0.031	3.722	0.032
28	3.652	0.035	3.798	0.031	3.957	0.032
29^{a}	2.545	0.043	3.178	0.033	3.434	0.036
31	3.435	0.038	3.590	0.033	3.897	0.036

Note: ^aItems for which the intercept was allowed to vary freely in sensitivity tests for our invariance models.