HHS Public Access

Author manuscript

Eur J Dev Psychol. Author manuscript; available in PMC 2021 March 12.

Published in final edited form as:

Eur J Dev Psychol. 2020 ; 17(5): 773–785. doi:10.1080/17405629.2020.1716724.

Validation of continuous measures of peer social interaction with self- and teacher-reports of friendship and social engagement

Robert L. Altman, **Brett Laursen**
Florida Atlantic University

Lynn K. Perry, **Daniel S. Messinger**
University of Miami

Abstract

The present study validates a new procedure that combines continuous measures of proximity (Ubisense) and vocalization (LENA) into measures of peer social interaction. The data were collected from 4 boys and 5 girls (ages 2–3 at the outset) on 8 separate days (3–4 hours per day) over the course of an academic year. Teacher reports of friendship were positively correlated with continuous measures of dyadic social interaction (i.e., the amount of time two children spent in proximity to one another, talking). Self-reports of reciprocated friendship were marginally correlated with continuous measures of dyadic social interaction, but only in the spring semester (when children were older and their reports of friendship more reliable). At the individual level, peer nominations of likeability, and teacher ratings of sociability and withdrawal were correlated with continuous measures of social interaction (i.e., the amount of time a child spent in proximity to other children, talking).

Keywords

Ubisense; LENA; Friendships; Social Engagement; Withdrawal

Some of the first empirical studies of child development involved collecting *in vivo* data to measure the affiliations of preschool children (Goodenough & Anderson, 1931). Little has changed in the intervening years. Expert observations remain the gold standard, underpinning the validity of child and teacher reports. It is easy to understand why scholars value expert observations of affiliation: They are both accurate and descriptive. New technologies are available that build on these advantages, providing continuous, simultaneous data on all children in a classroom. The goal of the present study is to validate a new procedure for assessing peer affiliation, one that integrates data obtained from continuous movement tracking (Ubisense: Killijian et al., 2016) and continuous vocalization

Address correspondence to laursen@fau.edu.

Data Availability Statement

The authors confirm that the data supporting the findings of this study are available within the article's supplementary materials.

recordings (LENA: Sangwan et al., 2015). We illustrate with data from an inclusive oral language preschool classroom for children with and without hearing loss.

Friendships matter, even to very young children. Preschool children with friends are more socially competent and accepted than children without friends (Vaughn, Colvin, Azira, Caya, & Kryzsik, 2001). Expert observations indicate that young children spend more time with those they reciprocally nominate as friends than with nonfriends, and more time with nonfriends than with classmates they do not like (Hartup et al., 1988; Masters & Furman, 1981). Similar findings emerge from teacher reports of friendship (Marshall & McCandless, 1957), which correlate with peer reports (Shin et al., 2014). Teacher and child reports also agree on the extent to which children are well liked by peers; expert observations corroborate these reports and suggest that well-liked children may enjoy status because they verbalize more during play than other children (Dunnington, 1957a, 1957b).

The primary drawback of expert observations is their impracticality: They are cumbersome, expensive, and time-consuming; typically, only one child can be observed at a time, for a small window of time. Child reports have their own challenges: Participants, especially young ones, are often uncooperative or distracted, and the data collection process is intrusive. Scholars are hesitant to deploy friendship nominations with children younger than 3 because they tend to be unreliable (Hartup, Glazer, & Charlesworth, 1967). Teacher reports are not typically used as a primary source of affiliation data because they can be confounded with perceptions of child characteristics and they may not be responsive to subtle shifts in patterns of affiliation. Still, given the challenge and expense of collecting expert observational data and the lack of available alternatives, most scholars are willing to tolerate the error inherent in child and teacher reports.

We focus on two technological advances that provide an alternative to expert observations. Ubisense, a commercially available radio frequency tracking system, can be used gauge affiliation through measures of proximity and mutual orientation (Messinger et al., 2019). LENA, a commercially available system of measuring child speech production, can be used to gauge social interaction through measures of expressive language (Perry et al., 2018). Using a different sample of children, we recently combined these two tools, demonstrating how they can be used to create a social contact score for each child in a preschool classroom (Perry et al., 2020). Here, we describe how these new forms of data converge with traditional child and teacher reports of peer affiliation.

Method

Participants

The participants included 9 children (4 boys, 5 girls) between the ages of 2 and 3 at the outset (M_{age}=30.33 months, SD=2.45). Participants were drawn from an English-dominant oral language inclusion classroom, located at a university-affiliated school in Miami, USA. Participants included 7 children who use hearing aids or cochlear implants and 2 typically hearing children. The participants included 2 African Americans, 3 Anglo Americans, 3 Hispanic Americans, and 1 child with a mixed ethnic background.

Procedure

All children in the class participated. Written consent was obtained from parents and from teachers (n=2 at the outset, then n=3 from the end of the fall semester onward). Peer nomination data were collected by a trained research assistant in a quiet classroom setting. Teachers completed questionnaires after class. Child and teacher reports of friendship were collected the same week as the continuous data, four times in the fall of 2017 (October-December) and three times in the spring of 2018 (January-March). Peer and teacher reports of social engagement were collected at the beginning and end of each semester. Continuous measures of social interaction were collected one day a week approximately every other week in the fall semester and approximately every third week in the spring semester. Each observation was 3–4 hours in length. Data were collected indoors, in a 7.56m × 4.78m classroom. Stability coefficients for each variable are given in terms of coefficient r, accompanied by confidence intervals and p-values (both one-tailed to reflect directional tests).

Measures

Continuous measures of social interaction (Ubisense and LENA).—Two objective measures of affiliation were integrated. Continuous measurements of child location were collected using the Ubisense Dimension4 system. Audio was collected using LENA Digital Language Processors. Each participant wore a vest (see Appendix) with two tags housed in the left and right rear pockets to provide information about orientation. LENA audio recorders were secured in the front pocket. The Ubisense system tracked each child's location to an accuracy of 15 cm using ultra-wide radio frequency identification. Audio files were analyzed using LENA Pro V3.4.0 pattern recognition software. We define social interaction as a vocalization made when a child was within 1.5 meters of another child, where both children were oriented (±45° degrees) toward one another. We created two measures of social interaction; z standardized to account for differences in the amount of time spent in class. The *dyadic continuous measure of social interaction* represents the number of vocalizations made when a child was in the presence of another child, divided by the amount of time the two children were in the classroom together (see Perry et al., 2020). A dyadic score was calculated for each pair of children in the classroom (N=36), by averaging the social interaction scores of one child when in the presence of the other child (e.g., the percentage of time child A vocalized in the presence of child B plus the percentage of time child B vocalized in the presence of child A, divided by 2). Scores were z-standardized within waves, then averaged across waves. Dyadic scores were stable from the beginning to the end of each semester (fall r=.71 [95% CI: .50, .84], p<.001; spring r=.55 [95% CI: .27, .74], p<.001), between assessment periods within semesters (fall range: r=.76 to .91 [M=.85]; spring range: r=.10 to .65 [M=.43]), and between the mean of the fall and the mean of the spring semesters (r=.82 [95% CI: .65, .91], p<.001). The *individual continuous measure of social interaction* represents the number of vocalizations made by a child in the presence of any other child in the classroom, divided by the amount of time the child was in class. An individual score was calculated for each child in the class (N=9). Scores were z-standardized within waves, then averaged across waves. Individual scores were stable from the beginning to the end of the fall semester (r=.91 [95% CI: .55, .98],

p=.001), between assessment periods within semesters (fall range: *r*=.89 to .95 [*M*=.92]; spring range: *r*=.12 to .77 [*M*=.45]), and between the mean of the fall semester and the mean of the spring semester (*r*=.89 [95% CI: .55, .98], *p*=.002). Stability coefficients from the beginning to the end of the spring semester were substantial, but only marginally significant (*r*=.58 [95% CI: −.28, .86], *p*=.10).

Friendship: Teacher reports.—Teachers identified all friendship dyads on a grid that listed each child in the class across the top and the side: "Friends are children who like one another and who play together." Teacher reports of friendship were stable from the beginning to the end of each semester (fall *r*=.32 [95% CI: −.01, .59], *p*=.03; spring *r*=.37 [95% CI: .05, .62], *p*=.01) between assessment periods within semesters (range *r*=.32 to .45 [*M*=.39]), and across the fall and spring semesters (*r*=.65 [95% CI: .31, .82], *p*=.001). In the fall semester, teachers identified an average of 5.0 friend dyads per wave of data collection (*SD*=1.0; min=4, max=6). In the spring semester, teachers identified an average of 4.7 friend dyads per wave of data collection (*SD*=2.5; min=2, max=8). Each dyad received a *teacher-report friendship score*, representing the number of times during a semester that two children were identified as friends.

Friendship: Child reports.—Children identified friends from an array of head and shoulder photos that included all 9 children in the class: "Who is your friend?" Unlimited same- and other-sex nominations were permitted. Reciprocated friendships were defined as dyads in which both partners nominated one another as friends. Reciprocated friendships were unstable from the beginning to the end of the fall semester (*r*=−.19 [95% CI: −.49, .15], *p*=.14), and between assessment intervals within the fall semester (range *r*=−.12 to .57 [*M*=−.14]), but were stable from the beginning to the end of the spring semester (*r*=.80 [95% CI: .64, .89], *p* <.001) between assessment intervals within the spring semester (range *r*=.17 to .25 [*M*=.21]), and across the fall and spring semesters (*r*=.48 [95% CI: −.02, .74], *p*=.03). In the fall semester, children reported an average of 7.5 reciprocated friendships per wave of data collection (*SD*=3.6; min=3, max=12). In the spring semester, children reported an average of 3.3 reciprocated friendships per wave of data collection (*SD*=1.3; min=2, max=5). Each dyad received a *child-report friendship score*, representing the number of times during a semester that the two children were reciprocated friends.

Social engagement: Teacher reports.—Teachers completed two 4-item measures from the *Child Behavior Scale* (Ladd & Profilet, 1996), rating each child's *sociability* (e.g., "Does this child play with other classmates enthusiastically?") and *withdrawal* (e.g., "Does this child prefer to play alone?"). Items were rated on a scale ranging from 1 (*Never*) to 3 (*Always*). Internal reliability was good (alpha=.87 to .88). Teacher ratings were stable from the beginning to the end of the fall semester (Sociability *r*=.91 [95% CI: .62, .98], *p*=.001; Withdrawal *r*=.92 [95% CI: .66, .98], *p*<.001), from the beginning to the end of the spring semester (Sociability *r*=.73 [95% CI: .13, .94], *p*=.01; Withdrawal *r*=.72 [95% CI: .11, .94], *p*=.01), and across the fall and spring semesters (Sociability *r*=.80 [95% CI: .29, .96], *p*=.01; Withdrawal *r*=.85 [95% CI: .43, .97], *p*=.001).

Social engagement: Peer report.—At the beginning and end of each semester, peer nominations of *likeability* ("Who do you like?") were obtained using a standard sociometric procedure. Children identified classmates from a board containing separate photos of each child in the class. Unlimited same- and other-sex nominations were permitted. In the fall semester, children nominated an average of 1.8 peers per wave of data collection (*SD*=0.2; min=1.7, max=2). In the spring semester, children nominated an average of 1.6 peers per wave of data collection (*SD*= 0.6; min=1, max= 2.2). Each participant received a score representing the number of nominations he or she received from classmates, divided by the total number of nominators. Peer reports of likeability were not stable from the beginning to the end of the fall semester (*r*=.07 [95% CI: −.62, .70], *p*=.42), nor across the fall and spring semesters (*r*=.30 [95% CI: −.45, .80], *p*=.21). Stability coefficients from the beginning to the end of the spring semester were substantial, but failed to reach conventional levels of statistical significance (*r*=.46 [95% CI: −.29, .86], *p*=.11).

Plan of Analysis

Missing data accounted for an average of 7.3% (*SD*=0.07; min=0%; max=17.1%) of reports for all variables included in the study. There was no attrition, but attendance varied. In the fall semester, children were present for either 3 or 4 of the observation sessions (*M*=3.8, *SD*=0.4). In the spring semester, children were present for either 2 or 3 of the observation sessions (*M*=2.9, *SD*=0.3). There were no missing teacher reports in either semester. Children participated in 3 or 4 (*M*=3.2, *SD*=1.2) peer nomination sessions in the fall semester and 2 or 3 (*M*=2.6, *SD*=1.0) peer nomination sessions in the spring semester. Missing data were imputed using the joint model imputation approach (McNeish, 2017).

Two sets of analyses were conducted to validate the continuous measures of social interaction. The first analyses (*N*=36 dyads) examined convergence between the dyadic continuous measure of social interaction and (a) teacher reports of friendship and (b) child reports of reciprocated friendship. We conducted two-way mixed, average measures, consistency intraclass correlations because the variables were designed to measure similar constructs and because the children participated in multiple, non-independent dyads (McGraw & Wong, 1996). The second analyses (*N*=9 children) examined convergence between the individual continuous measure of social interaction and (a) average teacher ratings of social engagement and (b) average peer ratings of social engagement. We selected Pearson's *r* interclass correlations because the variables did not represent identical constructs and because each child was included only once in the data. Convergence coefficients are accompanied by confidence intervals and *p*-values (both one-tailed to reflect directional tests). Analyses were conducted with SPSS (Version 25.0, IBM Corp., 2017).

Results

Table 1 presents results for dyadic measures of affiliation (*N*= 36 dyads). The dyadic continuous measure of social interaction was positively correlated with teacher reports of friendship in the fall and (at borderline levels of statistical significance) in the spring. The dyadic continuous measure of social interaction was correlated (at borderline levels of statistical significance) with child reports of friendship in the spring but not in the fall. Thus,

dyads with higher continuous social interaction scores tended to be rated as friends by teachers and children, although some of these associations fell short of conventional statistical significance. Teacher and child reports of friendship were positively correlated in the fall and spring.

Table 2 presents results for individual measures of affiliation (N= 9 children). The individual continuous measure of social interaction was positively correlated with child reports of likeability (spring only) and with teacher reports of sociability (fall and spring); the individual continuous measure of social interaction was negatively correlated with teacher reports of withdrawal (fall and spring). Thus, individuals with higher continuous social interaction scores were more apt to score higher on sociability and likeability, and score lower on withdrawal. Teacher reports of withdrawal and sociability were negatively correlated (fall and spring); neither were correlated with child reports of likeability (with the exception of a borderline statistically significant negative correlation with withdrawal in the spring).

Discussion

In this study, we paired Ubisense proximity measures with LENA vocalization recordings to create continuous measures of peer affiliation that assay dyadic friendship and individual sociability. The continuous measures of social interaction largely aligned with traditional indicators of peer affiliation, accounting for 24–50% of the variance in teacher reports of dyadic friendship and 49–62% of the variance in teacher reports of individual social engagement. Continuous measures of social interaction were not aligned with child reports of affiliation in the fall, but in the spring they accounted for 34% of the variance in child reports of dyadic friendship and 29% of the variance in child reports of individual likability.

Continuous measures of social interaction hold important advantages over traditional measures of friendship and sociability: they are unobtrusive, they simultaneously assess all children in a classroom, and they have fewer sources of error. Child reports, in particular, were plagued by low stability, especially in the fall semester when some of the participants were only 2 years old, which is at the low end of successful administration of the picture sociometric task (Hymel, 1983). Some might object to the fact that continuous measures of social interaction operationalize friendship and social engagement in purely objective terms, and they would have a point; continuous measures fail to capture the affective dimension contained in questions about friendship. Of course, much depends on how friendship is construed. Young children place a premium on companionship in friendships (Laursen & Hartup, 2002), suggesting that objective measures conceptually align with self-reports during this age period.

Scholars considering replacing expert observers with continuous measures of social interaction may be concerned about the expense. LENA and Ubisense carry considerable set-up costs (several thousand USD at the time of publication) that have no parallel in expert observations. But the more data that one collects, the more the cost differential narrows, because expert observations are more labor-intensive than continuous measures of social interaction, during data collection and afterwards. Note also that the cost of the equipment

should decrease (as is often the case with new technology), whereas the cost of the labor required for expert observations is apt to increase. Aside from costs, we see few other drawbacks to replacing expert observers and teacher reports with continuous measures of social interaction. Observer reports contain error, on the part of the observer and on the part of the coder. The technology described herein is not error-free; children are not always paying attention when others in close proximity speak to them, but expert observers are not immune to these errors either and are inclined to make other errors. Expert observations are also limited in terms of the amount of data that can be collected. Teacher reports of companionship, although inexpensive to obtain, carry few advantages over technological assessments. It is possible to deploy this technology in larger classrooms with more students, where information demands are apt to overload teacher abilities to accurately report how much time each child spends in the company of each other child, and confirmation bias (recalling answers from previous reporting periods) may give the illusion that teachers are reliable when in fact they may be missing subtle changes in social interaction patterns. Of course, if the researcher's goal is simply to measure the friendships of very young children, it may be difficult to justify the expense of continuous measurements. In these circumstances, our results should serve to remind scholars that child reports are less stable than teacher reports and that the two only weakly align; it may be prudent to collect information from both sources because they may capture somewhat different constructs.

Should continuous measures of child interaction replace child reports of friendship and likeability? We think not, because peer nominations reflect attitudes and affection, components of relationships and status that cannot be captured through the technologies described here.

Our study is not without limitations. Although the teacher and child reports provide converging evidence of the validity of the continuous measures of interaction, our case would be stronger had we also included observational data. It is important to note that sensor data were only collected in the classroom; we would expect greater convergence with teacher and peer reports had we also collected playground data. It is essential to replicate the current results with larger samples, which will produce more stable parameter estimates and improve the overlap between objective measurement, teacher ratings, and child reports. Small samples limit the generalizability of the results, as does our use of an oral language inclusion classroom. It is worth noting, however, that these limitations should increase error variance, thus decreasing convergence across measures. More classrooms with more verbally proficient students could yield greater convergence. The majority of communication in the classroom was in English, but not all participants came from families who spoke English at home. This may have contributed to the low reliability of the sociometric task, particularly in the fall. Typically, intraclass correlations are not negative. The (nonsignificant) negative value we obtained with child-reports of friendship implies an excess of within-group error variability. Finally, it is possible that some speech made in the proximity of a classmate was not heard or attended to by that classmate or was not directed to that classmate. As a consequence, reciprocated interaction may have been over-estimated, especially among the highly verbal. By the same token, verbosity might increase the chances that the socially central are erroneously identified as friends; it should not, however, adversely impact assessments of sociability or likeability.

Empirical assessments of peer affiliation have changed little in the past century, despite their well-documented limitations. The opportunity for a technological upgrade is at hand. Both friendship and sociability can be measured objectively with new tools that continuously assess location and speech. As these tools become more affordable and widespread, scholars of peer relationships should not hesitate to adapt them to new purposes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The project was funded by a grant to Daniel Messinger and Brett Laursen from the US National Science Foundation (IBSS-L-1620294), with additional support to Daniel Messinger from the Institute of Education Sciences (R324A180203) and to Brett Laursen from the US National Institute of Child Health and Human Development (HD096457).

Appendix

Left: Child wearing vest outfitted with LENA recorder (front pocket). Right: Backside of vest with pockets for Ubisense tags (left and right).

References

Dunnington MJ (1957a). Behavioral differences of sociometric status groups in a nursery school. Child Development, 28(1), 103–111. doi: 10.2307/1126005. [PubMed: 13404662]

Dunnington MJ (1957b). Investigation of areas of disagreement in sociometric measurement of preschool children. Child Development, 28(1), 93–102. doi: 10.2307/1126004 [PubMed: 13404661]

Goodenough FL, & Anderson JE (1931). Experimental child study. New York: Century.

Hartup WW, Glazer JA, & Charlesworth R (1967). Peer reinforcement and sociometric status. Child Development, 38(4), 1017–1024. doi: 10.2307/1127099. [PubMed: 4867866]

Hartup WW, Laursen B, Stewart MI, & Eastenson A (1988). Conflict and the friendship relations of young children. Child Development, 59(6), 1590–1600. doi:10.2307/1130673. [PubMed: 3208570]

Hymel S (1983). Preschool children's peer relations: Issues in sociometric assessment. Merrill-Palmer Quarterly, 29(3), 237–260. doi: https://www.jstor.org/stable/23086261.

Corp IBM. Released 2017. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.

Killijian M-O, Pasqua R, Roy M, Trédan G, & Zanon C (2016). Souk: Spatial observation of human kinetics. Computer Networks, 111(Supplement C), 109–119. doi: 10.1016/j.comnet.2016.08.008.

Ladd GW, & Profilet SM (1996). The Child Behavior Scale: A teacher-report measure of young children's aggressive, withdrawn, and prosocial behaviors. Developmental Psychology, 32(6), 1008–1024. 10.1037/0012-1649.32.6.1008.

Laursen B, & Hartup WW (2002). The origins of reciprocity and social exchange in friendships. New Directions for Child and Adolescent Development, 95, 27–40.

Masters JC, & Furman W (1981). Popularity, individual friendship selection, and specific peer interaction among children. Developmental Psychology, 17(3), 344–350. 10.1037/0012-1649.17.3.344.

McCandless BR, & Marshall HR (1957). A picture sociometric technique for preschool children and its relation to teacher judgments of friendship. Child Development, 28(2), 139–147. doi:10.2307/1125876. [PubMed: 13427066]

McGraw KO, & Wong SP (1996). Forming inferences about some intraclass correlation coefficients. Psychological Methods, 1(1), 30–46. doi:10.1037/1082-989X.1.1.30.

McNeish D (2017). Missing data methods for arbitrary missingness with small samples. Journal of Applied Statistics, 44, 24–39. doi: 10.1080/02664763.2016.1158246.

Messinger DS, Prince EB, Zheng M, Martin K, Mitsven SG, Huang S, Stölzel T, Johnson N, Rudolph U, Perry LK, & Laursen B (2019). Continuous measurement of dynamic classroom social interactions. International Journal of Behavioral Development, 43(3), 263–270. doi: 10.1177/0165025418820708.

Perry LK, Mitsven SG, Custode S, Vitale L, Laursen B, Song C, & Messinger DS (2020) Preschool dynamics: Reciprocal patterns of peer speech. Manuscript submitted for publication.

Perry LK, Prince EB, Valtierra AM, Rivero-Fernandez C, Ullery MA, Katz LF, Laursen B, & Messinger DS (2018). A year in words: The dynamics and consequences of language experiences in an intervention classroom. PloS ONE, 13(7): e0199893. doi: 10.1371/journal.pone.0199893. [PubMed: 29979740]

Sangwan A, Hansen JHL, Irvin DW, Crutchfield S, & Greenwood CR (2015). Studying the relationship between physical and language environments of children: Who's speaking to whom and where? 2015 IEEE Signal Processing and Signal Processing Education Workshop (SP/SPE), 49–54. 10.1109/DSP-SPE.2015.7369526.

Shin N, Kim M, Goetz S, & Vaughn BE (2014). Dyadic analyses of preschool-aged children's friendships: Convergence and differences between friendship classifications from peer sociometric data and teacher's reports. Social Development, 23(1), 178–195. 10.1111/sode.12043.

Vaughn BE, Colvin TN, Azria MR, Caya L, & Krzysik L (2001). Dyadic analyses of friendship in a sample of preschool-age children attending Head Start: Correspondence between measures and implications for social competence. Child Development, 72(3), 862–878. 10.1111/1467-8624.00320. [PubMed: 11405587]

Table 1.

Convergence between Dyadic Measures of Peer Affiliation: Within-Dyad Intraclass Correlations

Variable	1 ([95% CI] *p*)	2 ([95% CI] *p*)	3 ([95% CI] *p*)	*M* (*SD*)
1. Dyadic continuous measure of social interaction		.34 ([−.16, ∞] .06)	.24 ([−.33, ∞] .10)	0.00 (0.61)
2. Child-reported Friendship	−.24 ([−1.0, ∞] .37)		.84 ([.71, ∞] .001)	0.28 (0.66)
3. Teacher-reported Friendship	.50 ([.12, ∞] .01)	.32 ([−.19, ∞] .06)		0.39 (0.73)
M (*SD*)	0.00 (0.93)	0.83 (0.88)	0.28 (0.57)	

Note. N=36 dyads. Fall 2017 correlations below the diagonal and spring 2018 correlations above the diagonal. Dyadic continuous measure scores represent the number of vocalizations one child made in the presence of another child, divided by the amount of time the two children were in the classroom together; for each dyad, scores were averaged across partners, z-standardized within waves, and averaged across waves. Child-report scores represent the number of times during a semester that two children were reciprocated friends. Teacher-report scores represent the number of times during a semester that teachers identified two children as friends. One sided confidence intervals (in brackets) and *p*-values are presented.

Table 2.

Convergence between Individual Measures of Peer Affiliation: Within-Individual Interclass Correlations

Variable	1 ([95% CI] *p*)	2 ([95% CI] *p*)	3 ([95% CI] *p*)	4 ([95% CI] *p*)	*M* (*SD*)
1. Continuous measure of social interaction		.54 ([−.07, ∞] .07)	.79 ([.38, ∞] .01)	−.63 ([−∞, −.07] .04)	0.00 (0.50)
2. Child-reported likeability	−.21 ([−.71, ∞] .30)		.56 ([−.04, ∞] .06)	−.55 ([−∞, .05] .06)	0.26 (0.09)
3. Teacher-reported sociability	.70 ([.19, ∞] .02)	−.15 ([−.68, ∞] .35)		−.73 ([−∞, −.25] .01)	2.79 (0.23)
4. Teacher-reported withdrawal	−.79 ([−∞, −.38] .01)	.10 ([−∞, .65] .40)	−.72 ([−∞, −.23] .01)		1.25 (0.17)
M (*SD*)	0.00 (0.60)	0.23 (0.10)	2.43 (0.48)	1.25 (0.35)	

Note. *N*=9 children. Fall 2017 correlations below the diagonal and spring 2018 correlations above the diagonal. Continuous measure scores were z standardized within waves and averaged across waves. Likeability represents the number of nominations a child received from classmates, divided by the number of nominators. Sociability and withdrawal were rated on a scale ranging from 1 (*never*) to 3 (*always*), averaged across reporters. One sided confidence intervals (in brackets) and *p*-values are presented.