Original Research Article

Assessment
1–16

DOI: 10.1177/10731911221083613
journals.sagepub.com/home/asm
SAGE

Ecological and Convergent Validity of Experimentally and Dynamically Assessed Capacity for Social Contingency Detection Using the Perceptual Crossing Experiment in Adolescence

Karlijn S. F. M. Hermans[1*], Olivia J. Kirtley[1], Zuzana Kasanova[1], Robin Achterhof[1], Noëmi Hagemann[1], Anu P. Hiekkaranta[1], Aleksandra Lecei[1], Leonardo Zapata-Fonseca[2,3], Ginette Lafit[1], Ruben Fossion[2], Tom Froese[4], and Inez Myin-Germeys[1]

Abstract

The Perceptual Crossing Experiment (PCE) captures the capacity for social contingency detection using real-time social interaction dynamics but has not been externally validated. We tested ecological and convergent validity of the PCE in a sample of 208 adolescents from the general population, aged 11 to 19 years. We expected associations between PCE performance and (a) quantity and quality of social interaction in daily life, using Experience Sampling Methodology (ESM; ecological validity) and (b) self-reported social skills using a questionnaire (convergent validity). We also expected PCE performance to better explain variance in ESM social measures than self-reported social skills. Multilevel analyses showed that only self-reported social skills were positively associated with social experience of company in daily life. These initial results do not support ecological and convergent validity of the PCE. However, fueled by novel insights regarding the complexity of capturing social dynamics, we identified promising methodological advances for future validation efforts.

Keywords

dynamic systems theory, perceptual crossing experiment, self-report, experience sampling methodology, adolescence, social skills

The development of social skills is highly important in adolescence (Blakemore & Mills, 2014). This period marks increased interaction with peers, constituting novel social environments to which adolescents must adapt. A key aspect of successful interpersonal functioning, therefore, is the capacity to flexibly adapt to novel social roles and environments (Dahl et al., 2018), reflecting social dynamics. Given that the foundations of adult interpersonal functioning are laid during adolescence, altered social development in adolescence could have long-lasting consequences if left unaddressed (Gresham, 2016; Zampella et al., 2020). Impaired interpersonal functioning is a key component of many mental disorders with their onset in late adolescence, such as depression and psychosis (Kessler, 2007). An impaired capacity to detect and respond to social contingencies, resulting in the impaired capacity to adapt to changing social environments, could therefore be viewed as a transdiagnostic factor in a range of mental disorders, recently also referred to as disorders of social interaction (e.g., Leong & Schilbach, 2019).

To date, the dynamic nature of the capacity for social contingency detection has generally not been an integral part of the assessment of impaired social capacity. The capacity to detect social stimuli and responding to those in real time requires two individuals. Arguably, therefore, the study of this capacity from an individual perspective rather than a dynamic perspective may account for the lack of robust associations between findings from laboratory experiments on social cognition and social interaction in the

[1]KU Leuven, Belgium
[2]National Autonomous University of Mexico, Mexico City
[3]University of Heidelberg, Germany
[4]Okinawa Institute of Science and Technology Graduate University, Japan

*Karlijn S. F. M. Hermans now affiliated to: Developmental and Educational Psychology, Faculty of Behavioral and Social Sciences, Leiden University, Leiden, The Netherlands; Department of Psychology, Education and Child studies, Erasmus School of Social and Behavioral Sciences, Erasmus University Rotterdam, Rotterdam, The Netherlands

Corresponding Author:
Inez Myin-Germeys, Center for Contextual Psychiatry, Department of Neurosciences, KU Leuven, ONV Herestraat 49, Box 1029, Leuven 3000, Belgium.
Email: inez.germeys@kuleuven.be

real world (Osborne-Crowley, 2020). If assessments of the capacity for social contingency detection are not associated with social behavior and experiences in daily life, this casts doubt upon the utility of these assessments for early detection and prevention of social impairments in adolescents. Therefore, including the real-life aspect of social interaction dynamics in the assessment of the capacity for social contingency detection could improve early detection of social alterations in adolescence.

Social behavior and experiences in daily life can be captured using Experience Sampling Methodology (ESM), which includes repeated, within-person assessments across fluctuating daily environments (Csikszentmihalyi & Larson, 2014; Myin-Germeys et al., 2018). Although ESM closely reflects individuals' real-world behavior and experiences in fluctuating contexts, two challenges arise when using ESM to inform early detection and prevention of altered social development. First, ESM captures individual reflections of the dynamic environments encountered during daily life, as opposed to social dynamics *per se* (e.g., Myin-Germeys, 2020). That is, in defining social interaction as the continuous adaptation to the other, the assessment of social dynamics requires real-time interaction. Although momentary, dyadic experience sampling is possible (Laurenceau & Bolger, 2012), ESM studies typically do not allow for the assessment of how social dynamics unfold in real time. Reporting about the interaction typically disrupts the social interaction at that moment and takes place seconds to hours after the interaction has been established, which rules out the assessment of social dynamics. Second, the replicability of ESM findings is arguably restricted due to the method's naturalistic nature. This diminishes comparability between individuals and, consequently, may hamper the identification of early social impairments at the group level. Despite these challenges, social interaction measures collected with ESM provide a reflection and granular assessment of social behavior and experience in daily life (Mote & Fulford, 2020; Myin-Germeys et al., 2018). Hence, if an assessment of the capacity for social contingency detection is associated with social behavior and experience captured with ESM, this assessment's ecological validity would be supported.

To include social dynamics in the assessment of the capacity for social contingency detection while maintaining experimental control, thereby contributing to replicability, Virtual Reality (VR) techniques reflecting a variation of interaction dynamics and graphical realism have shown promise for the study of social interaction (Pan & Hamilton, 2018). As capturing social interaction dynamics is highly time sensitive, and can easily break down due to small delays in responding in real time, VR techniques for studying minimal interactions have been designed (Auvray et al., 2009; Pan & Hamilton, 2018). The Perceptual Crossing Experiment (PCE; Auvray et al., 2009; Auvray & Rohde, 2012; Lenay & Stewart, 2012) is a type of minimalistic VR, which is high in interaction dynamics and low in graphical realism (Figure 2; Pan & Hamilton, 2018). The PCE allows pairs of participants to establish an interaction only relying on detecting and responding to each other's, hence social, stimuli. Humans are very sensitive to social cues. They continuously need to detect (moving) stimuli as being social to respond to those and establish an interaction by coordinating their behavior with their interaction partner. This phenomenon has previously been demonstrated in infants, using the double TV paradigm (Trevarthen & Aitken, 2001). Infants were shown a videoscreen, which either displayed their mother responding to them in real-time or a videotape of an earlier interaction. The infants became distressed if their interaction with the mother did not include a real-time response component (i.e., it was delayed). Similarly, the PCE examines how individuals detect the pattern of interaction as being social. The PCE examines the interaction between two individuals on a behavioral, sensorimotor level, in line with an embodied approach to cognition (Newen et al., 2018). Some scholars even suggest that this behavioral coordination to each other—based on social contingency detection—is a prerequisite for developing social cognition (De Jaegher et al., 2010).

In isolating this most basic component of social interaction (e.g., Trevarthen & Aitken, 2001), the PCE relies only on tactile, sensorimotor feedback. The focus on the tactile modality in the PCE, rather than verbal interaction, stems from the first developed modality for interaction in infants and may, therefore, provide the most simple, minimalistic form of interaction.

Specifically, the PCE requires pairs of participants to interact with each other in a one-dimensional virtual environment (i.e., they cannot see or hear each other), in which they are embodied as avatars. Participants are asked to establish an interaction with each other, relying only on a vibration on their hand, which occurs when the avatars cross each other (Auvray et al., 2009; Froese, Iizuka, & Ikegami, 2014). Within the virtual environment, there is also a nonresponsive and moving, and a nonresponsive static entity, providing the same tactile feedback as the avatar when crossed. The only way to detect the other (i.e., detecting social contingency) is through interacting, thus responding to the other using sensorimotor feedback (i.e., moving based on the vibrations). Participants engage in 1-minute rounds in which they are instructed to cooperate. They are not given any feedback on whether they are interacting with (i.e., crossing) their partner or with other entities, such that their interaction remains hidden ("implicit") to them. After each round, participants are asked to report on their awareness of the interaction established during that round. Within the PCE assessment, we distinguish three components of the capacity for social contingency detection. After each round, we capture individuals' *awareness*

of their "implicit" social *behavior* (operationalized as *time spent together*), in addition to their explicit judgment of this behavior expressed by correctly detecting the other participant—*accuracy*—during the implicit interaction.[1] To date, however, the assessment of social dynamics to capture the capacity for social contingency detection using minimalistic VR is novel in adolescence.

The most frequently used measures of the capacity for social contingency detection, especially in adolescence, are self-reported social skills questionnaires (Gresham, 2016; Little et al., 2017). These are completed by adolescents themselves or by their caregivers or teachers, and usually include reflections on adolescents' average social behavior in predefined situations. These questionnaires, however, frequently ignore variation across contexts and interaction partners. For example, asking how adolescents behave toward their classmates ignores variation (i.e., dynamics) between different classmates across different situations (Rose & Fischer, 2008), which again hampers the assessment of real-time social dynamics. The advantage of a questionnaire-based self-report is the comparability among individuals, as the items and predefined situations are the same for everyone. However, this limits the ecological validity of questionnaires for social functioning, as demonstrated by a study finding weak associations between the Social Functioning Scale (Birchwood et al., 1990) and the quality and quantity of real-world social interactions, using ESM (Schneider et al., 2017). Self-reports of social skills are also time dependent, show low agreement between different raters, and focus on the judgment of the behavior rather than the behavior itself (Gresham, 2016). However, they have also shown to relate to adjustment difficulties and learning achievements (Del Prette et al., 2012; Gresham, 2016), as well as psychopathological symptoms (Van der Ploeg & Scholte, 2013). In addition, despite their shortcomings, this type of questionnaire is frequently used and at least provides a good reflection of the average self-reported social skills level. Therefore, as a measure of social skills, an association between PCE capacity for social contingency detection and self-reported social skills would support the convergent validity of the PCE.

Previous small-scale studies in adults using the PCE have demonstrated that participants can adequately establish a social interaction within the minimalistic virtual environment. This was supported by the association between heightened awareness of the interaction and mutual correct detection of the other, compared with single detection of the other and no detected interaction (Froese, Iizuka, & Ikegami, 2014; Froese et al., 2020; Zapata-Fonseca et al., 2016). We replicated this proof-of-principle of the PCE in the first study in adolescents, demonstrating similar findings with a shortened paradigm (Hermans et al., 2020). We also provided evidence for the developmental course of the capacity for social contingency detection throughout adolescence, as evidenced by late adolescents' faster behavioral social coordination during the experiment, in addition to a higher average level of performance in this group compared with younger adolescents (Hermans et al., 2021). These findings provided the first evidence for the developmental course of the capacity for social contingency detection across adolescence. As these first PCE findings in adolescence and their similarity to adult findings may be interpreted as content validity of the PCE to capture the capacity for social contingency detection, the next validation steps involve associations with external measures of social skills and behaviors (Cronbach & Meehl, 1955).

Current Study

The current study will investigate the ecological and convergent validity of the PCE, contributing to the improved assessment of the capacity for social contingency detection in adolescence. In addition, we will study whether the PCE has higher ecological validity than a self-reported social skills questionnaire, as the former includes a dynamic component of interaction, while the latter does not. We will test associations between PCE capacity for social contingency detection (i.e., awareness, accuracy, and time spent together) and (a) social interaction (i.e., probability and experience of company) in daily life, assessed with ESM to evaluate ecological validity, and (b) self-reported social skills assessed with the Social Skills Questionnaire (Vragenlijst Psychosociale Vaardigheden [VPV]; Van der Ploeg & Scholte, 2013) to evaluate convergent validity. To test which measure of the capacity for social contingency detection is most strongly associated with daily life social interaction, (c) the association between self-reported social skills and social interaction in daily life will be studied, in addition to (d) comparing PCE capacity for social contingency detection and self-reported social skills as predictors with social interaction in daily life as outcome. We hypothesize the following:

Hypothesis 1 (H1): We expect positive associations between PCE capacity for social contingency detection and social interaction in daily life.
Hypothesis 2 (H2): We expect positive associations between PCE capacity for social contingency detection and self-reported social skills.
Hypothesis 3 (H3): We expect positive associations between self-reported social skills and social interaction in daily life.
Hypothesis 4 (H4): We expect a stronger association between PCE capacity for social contingency detection and social interaction in daily life compared with self-reported social skills and daily life.

Method

Sample and General Procedure

Participants were recruited in secondary schools, as part of the "SIGMA project": a large, accelerated longitudinal study of adolescent well-being in the general population (for full details of the procedure and sample, see Kirtley et al., 2021, April 2). The sample was recruited in the first, third, and fifth year of secondary school, covering the age range between 11 and 19 years. We clustered adolescents per school year, reflecting their time and experience in secondary education and, therefore, reflecting developmental stages (e.g., Blakemore & Mills, 2014). These groups are hereafter referred to as early, mid, and late adolescence, respectively.[2]

Participation in the SIGMA study required a 100-minute classroom session including the completion of the PCE and the VPV (as part of a broader test battery), followed by six consecutive days during which participants completed brief questionnaires on a dedicated device (ESM). During classroom sessions where it was practically feasible, eight participants per classroom were randomly selected to perform the PCE. Random selection of participant numbers was done before the classroom session started, using a randomization website (random.org). After full participation, participants received a 10 Euro voucher for an online store. Ethical approval for the SIGMA study was obtained from the local Medical Ethics Committee (Ref: S6 1395). This study was post-registered (a form of pre-registration that occurs following data collection, but before conducting the analyses; Benning et al., 2019) on the Open Science Framework website (https://osf.io/bxzpk).

PCE

The procedure and full details of the PCE in adolescents have been fully described elsewhere (Hermans et al., 2020). In brief, participants were randomly assigned to pairs of players (i.e., dyads) and instructed to move their avatar, using a trackball with their dominant hand, through a shared, loop virtual one-dimensional space (Figure 1). Their collaborative task was to find each other in this space by interacting, using only tactile feedback they received as a vibration on their hand each time they crossed each other. They could not hear or see each other, or their avatar in virtual space, and hence could not communicate in any other way than moving the trackball back and forth, guided by the vibrations. The experiment consisted of six 1-minute rounds with random starting positions of both players' avatars. Within the virtual environment, participants could encounter the other player's avatar (responsive and moving), the other player's shadow (nonresponsive and moving), and a static object (nonresponsive and not moving). Participants were instructed to click a button with their other hand at the moment they were most confident that they were crossing the other, that is when they believed that the vibration was induced by the other player. To successfully identify the other from the moving shadow, the participant must interact with their partner, as the shadow is also moving but not responsive. Participants did not receive any feedback on their or their partner's performance during or after the experiment.

Three variables assess basic capacity for social contingency detection: awareness of the interaction (awareness), the proportion of correct clicks (accuracy), and time in seconds per round spent together with the other avatar (time spent together). We computed the PCE variables awareness, accuracy, and time spent together as aggregated scores per individual across six rounds. Details on the computation of PCE variables are provided elsewhere (Hermans et al., 2020).

Awareness. After each 1-minute round, participants rated three items on a 7-point Likert-type scale ranging from 1 "*not at all*" to 7 "*very much*": "To what extent did you feel that the other could sense your presence?," "To what extent did you feel you were doing something together?," and "How confident were you that you clicked correctly?" The variable awareness was calculated as the mean of these three items as they loaded onto a single factor (Hermans et al., 2020). If participants did not click within a round, awareness was computed as the mean of the first two items only. Cronbach's alpha of cases in which all three items were completed ranged from .85 to .91 across the six rounds. The correlation between items in which case two items were completed (i.e., no click) ranged from .61 to .95 across the six rounds. The exact values per round and missing values for the confidence item have been listed in Table S1A.

Accuracy. The proportion of correct clicks was computed as the number of correct clicks over six rounds, divided by the participant's total number of clicks (maximum six).

Time Spent Together. The time per round that participants spent together with the other player was computed as seconds of the total minute during which participants were in close proximity of the other (same value for both participants within a dyad).

Familiarity With the Other. After the random assignment of dyads and before the experiment started, participants were asked to complete two items on familiarity with the other. These captured the frequency of contact with the partner within a dyad ("How often do you interact/engage with this classmate?"), rated from 1 "never" to 7 "every day for most of the day," and the quality of contact with this partner ("How well do you get along with this

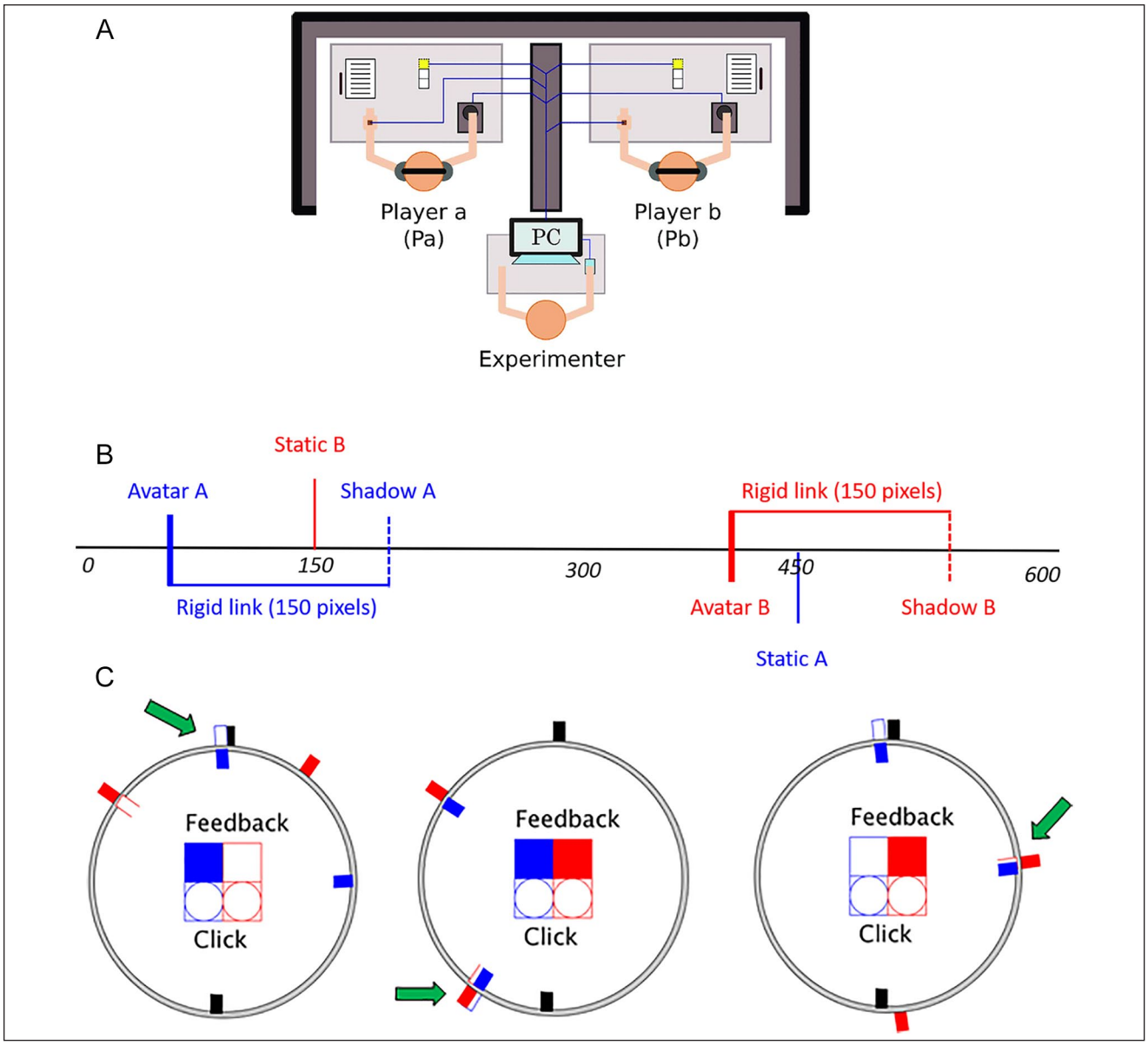

Figure 1. A: Set-Up PCE, Reproduced From Froese, Iizuka, and Ikegami (2014). B: The Virtual One-Dimensional Space With Locations of and Relations Between Entities (Invisible to Participants); Avatar A and B Representing Player A and B; Shadow A and B Representing the Shadow of Avatar A and B; and Each Player's Static Object. C: Loop One-Dimensional Virtual Space Showing Crossing With the Static Object, Crossing With the Other Avatar, Crossing With the Other Shadow, From Left to Right. The Colored Box in the Middle Shows the Vibration (Feedback) Perceived by Player A and B

classmate?"), rated from 1 "not well at all" to 7 "very well." Based on a significant and high correlation between these two items ($\rho = .72$, $p < .001$), the person-level mean was computed and used as a priori variable "familiarity" in the analyses to control for familiarity with the other player within a dyad (as preregistered). This addresses the potential variation explained by nestedness within class groups, as participants within the baseline session were part of different class groups within the same school.

Psychosocial Skills Questionnaire

The Psychosocial Skills Questionnaire (VPV; Van der Ploeg & Scholte, 2013) consisted of 36 statements reflecting self-reported inter- and intrapersonal skills. The response scale for each statement ranged from 1 "completely disagree" to 5 "completely agree." The interpersonal subscale includes subscales of relational skills (e.g., "Actively contact peers") and affective skills (e.g., "Take others' feelings into account"). The intrapersonal score consists of the subscales

of self-guidance (e.g., "Always complete [home]work entirely") and self-awareness (e.g., "Be open to criticism and comments"). VPV total score was computed as the sum score of 36 items per individual. The internal consistency of the total score of social skills (sum score of all items) has been reported as Cronbach's alpha = .93 (Van der Ploeg & Scholte, 2013). The item "I never let a difference of opinion get easily out of hand" was reworded to "I let a difference of opinion get easily out of hand," based on reported difficulty with the double negative in this statement. The questionnaire has shown good reliability and sufficient construct validity (Van der Ploeg & Scholte, 2013). The internal consistency for VPV total score in our sample was Cronbach's alpha = .89, and only this score was used in the current study, referred to as self-reported social skills.

ESM

We collected repeated self-reports in daily life using a mobile application (mobileQ; Meers et al., 2020) on a dedicated device loaned to the participants (Motorola Moto E4 Android smartphone). This device notified participants 10 times per day for six consecutive days to complete 3-minute brief questionnaires with maximum of 47 items (questionnaires hereafter referred to as beeps). A semi-random time sampling scheme was used in which notifications occurred at random time points within 90-minute intervals, with at least 15 minutes between consecutive notifications. The sampling took place between 8:30 a.m. and 10:30 p.m. More details of the ESM procedure are described in the SIGMA study protocol (Kirtley et al., In preparation). ESM self-reported questionnaires included a range of items, all publicly available via the ESM Item Repository (www.esmitemrepository.com; Kirtley et al., 2019). For the current study, only ESM items on social company and experiences of company were used: hereafter referred to as company and social experience.

Company. Participants indicated whether they were alone or in company at the moment of the beep. They could select multiple responses to specify their company, such as "parents" and "friends." For the current study, the variable "company" was used as a binary variable indicating whether participants indicated being in company or alone at the moment of the beep.

Social Experience. If participants were in company, they completed follow-up items regarding their experiences of this company. The variable "social experience"[3] comprised the mean of three items: "I feel at ease in this company," "I feel appreciated in this company," and "I belong [in this company]." The within- and between-person reliability of these items was good, with a Cronbach's alpha of social experience of .90 and .95, respectively.

Statistical Analysis

We performed statistical analyses using STATA 14.2 (StataCorp., 2015). ESM data have a hierarchical structure. These data were analyzed using multilevel mixed-effects regression models, assuming a three-level structure (repeated beeps nested within individuals within schools). For each analysis, age group (three levels: early, mid, and late adolescence) and familiarity with the other player were added as covariates. Analyses including PCE performance have been separately conducted for each of the three PCE variables (awareness, accuracy, and time spent together) as we aim to test associations with daily life for each conceptually different element of the PCE. For testing the association between independent variable X and dependent variable Y, the model for the i-th participant, in the j-th school at time t is the following: Y_{tij} = gamma_{0ij} + b1 × Age_i + b2 × VPV_i + b3 × Age × VPV_i + b4 × Familiarity_i + e_tij (Level 1); gamma_{0ij} = b00i + v_{ij} (Level 2); b00j = b000 + u_j (Level 3).[4] We report the results both uncorrected and corrected for multiple comparisons using the Bonferroni correction: $\alpha = .05/3$ (for the three separate PCE variables) = .017. For each association, a random intercept was allowed for school, accounting for potential differences between schools. For the analyses with PCE variables as the outcome, this was extended with a random intercept for dyad nested within school, and for the analyses with ESM variables as the outcome; this was extended with a random intercept for individual nested within school. In the tables, the p values that are significant after Bonferroni correction are highlighted in bold and the p values that would be significant without Bonferroni correction are highlighted in italics.

Primary Analyses. To test the association between PCE capacity for social contingency detection and social interaction in daily life (H1), we performed separate analyses with PCE variables awareness, accuracy, and time spent together as independent variables and ESM variables company (logistic) and social experience (linear) as dependent variables. To test the association between PCE capacity for social contingency detection and VPV self-reported social skills (H2), we performed separate multilevel linear mixed-effect regression analyses with PCE awareness, accuracy, and time spent together as independent variables and VPV self-reported social skills as the dependent variable. To test the association between self-reported social skills and social interaction in daily life (H3),[5] we used self-reported social skills as the independent variable and company (logistic) and social experience (linear) as dependent variables. Finally, to compare VPV self-reported social skills and PCE capacity for social contingency detection in explaining variance in social interaction in daily life (H4), we added VPV self-reported social skills to each of the PCE variables in the

same model as independent variables and company and social experience separately as dependent variables, resulting in six analyses. We standardized the regression coefficients of the independent variables if applicable. This was computed by the following formula for standardizing coefficients in mixed model effects: beta_standardized = (beta_non-standardized × *SD*(predictor))/*SD*(outcome) (Hox et al., 2010).

Exploratory Analyses. We conducted two exploratory analyses to improve our understanding of the results. These were not preregistered and were therefore added to the transparent changes document (https://osf.io/nhs6d/?view_only=4f7ea37b3cf648bdb7da48b1ce29dbb5). First, to test whether the associations differed across age groups, we added the interaction between age group and the independent variable after each analysis. We used the "margins" command to obtain the interaction coefficients for each age group. Second, to test whether interacting with company in daily life provided different findings compared with just being in company without interacting, we repeated the analyses with ESM social interaction as the outcome (company and social experience) including only the beeps in which participants reported to be interacting with their company. If participants were with others, they also rated the extent to which they interacted with their company ("We are doing something together"), on a scale from 1 "not at all" to 7 "very much." An interaction score of at least 2 indicated at least some level of interaction, and this was used to include the beeps at moments when participants were interacting with their company. The absence of interaction (interaction score of 1) could, for instance, be reported when participants were reading in the same room with others.

Sample Size Rationale. As the PCE was included as part of a larger study (Kirtley et al., in preparation), we maximized the number of participants as far as this was practically feasible. While a power calculation was conducted for the larger study, we did not conduct a separate power analysis for the PCE sample. We were unable to perform a post hoc sensitivity power analysis as there were no sufficiently powered or comparable previous studies available from which to draw parameter estimates (Anderson et al., 2017). Given the nature of PCE and ESM data—within-person repeated measures—the current sample size exceeds the typical average sample size in similar designs (e.g., intensive longitudinal studies; Lafit et al., 2021).

Results

Sample and Data Characteristics

The initial sample consisted of 208 participants who completed the PCE, including 80 early adolescents, 48 mid adolescents, and 80 late adolescents. For three participants, "awareness" data were missing, so the final sample of participants who completed the PCE comprised 205 adolescents. For 42 participants, no complete VPV data were available. According to the VPV manual (Van der Ploeg & Scholte, 2013), the total score can only be calculated if all items are completed, as this comprises the sum score of all 36 items. We report on a sample size of 166 for testing H2 and H4, excluding these participants.[6] For the analyses using only company and social experience in interaction, 13.9% of the entire number of completed beeps were excluded as these reflected company without interaction. For three participants, there were missing data on demographics and the PCE awareness variable. For 23 participants, no ESM data were available. We report on a sample size of 185 for testing H1. The sample size for testing H3 included 150 participants who completed all three measures. Table 1 shows demographics and other sample details. Late adolescents completed fewer daily questionnaires compared with both younger age groups. As part of the reporting guidelines we followed for ESM research reporting in adolescence (van Roekel et al., 2019), we have added information on valid beeps per age group (Table 1).

Findings Primary Analyses

H1: Association Between PCE Social Capacity and Social Interaction in Daily Life (ESM). The associations between PCE capacity for social contingency detection (awareness, accuracy, time spent together) and social interaction (company and social experience) in daily life are shown in Table 2.

We found no statistically significant association between PCE awareness, accuracy, and time spent together and being in company in daily life. In all three analyses, there was a significant association between age group and being in company, such that, on average and compared with early adolescents, older adolescents were less likely to report being in company in daily life.

We found no statistically significant association between PCE accuracy/time spent together and social experience in daily life. In these two separate analyses, there was a significant association between the mid-adolescent age group and social experience in daily life, such that, on average and compared with early adolescents, this age group reported a lower social experience of company in daily life. There was a statistically significant, positive association between PCE awareness and social experience in daily life, such that the higher the reported awareness of PCE interaction, the higher the quality of the interaction reported in daily life. These effects fell short of significance after Bonferroni correction ($\alpha = .05/3 = .017$).

H2: Association Between PCE Social Capacity and Self-Reported Social Skills (VPV). We found no statistically significant

Table 1. Sample Characteristics of the Total Number of Participants Who Completed the PCE; Aggregated Mean Scores and Standard Deviations for VPV, ESM, and PCE Variables.

	Total sample $N = 208$[a]	Early adolescence[b]—Year 1 $n = 79$	Mid adolescence[b]—Year 3 $n = 46$	Late adolescence[b]—Year 5 $n = 80$
Age (years)	14.4 (2.02)	12.2 (0.55)	14.6 (0.80)	16.5 (2.02)
Range	11–19	11–14	13–17	16–19
Gender				
male/female in %	40/60	44/56	26/74	45/55
VPV ($N = 166$)	134.0 (14.3)	138.0 (12.6)	131.3 (18.6)	130.6 (11.3)
Range	86–167	106–167	86–162	101–160
ESM ($N = 185$)[c]				
Company *in %*	84.7	90.9	80.3	79.0
Social experience	5.8 (1.5)	6.0 (1.5)	5.7 (1.5)	5.7 (1.4)
PCE ($N = 208$) [a]				
Awareness ($N = 205$)	3.7 (1.2)	4.0 (1.2)	3.6 (1.0)	3.4 (1.3)
Accuracy	0.4 (0.3)	0.3 (0.3)	0.4 (0.2)	0.5 (0.3)
Time spent together	20.0 (0.3)	18.5 (0.2)	19.5 (0.5)	22.0 (0.4)
Number valid beeps	30.6 (11.4)	31.8 (11.1)	31.4 (10.3)	27.9 (12.5)

Note. VPV = vragenlijst psychosociale vaardigheden; ESM = experience sampling methodology; PCE = perceptual crossing experiment.
[a]Missing data on demographics and awareness (PCE) for $n = 3$. Awareness: mean score of 2 or 3 items, rated on 7-point Likert-type scale. Accuracy: proportion correct clicks out of total number of clicks. Time spent together: back-transformed in seconds of total round (60 seconds). [b] The minimum and maximum age per age group overlaps across age groups as these are based on school year (1, 3, 5) within secondary education. [c] Percentage in company versus alone was computed across all beeps.

Table 2. Relationship Between PCE Variables and Daily Life Social Interaction (ESM).

	ESM					
	Company			Social experience		
	B (SE)	95% CI	*p*	*B (SE)*	95% CI	*p*
PCE awareness	0.08 (0.08)	[−0.09, 0.24]	.359	0.14 (0.06)	[0.01, 0.26]	*.029*
Year (age group)						
5 vs. 1	−0.92 (0.28)	[−1.47, −0.37]	**.001**	−0.34 (0.19)	[−0.72, 0.03]	.071
5 vs. 1	−0.97 (0.29)	[−1.54, −0.40]	**.001**	−0.09 (0.17)	[−0.44, 0.25]	.586
5 vs. 3	−0.05 (0.26)	[−0.56, 0.46]	.851	0.25 (0.20)	[−0.14, 0.64]	.214
Familiarity	−0.01 (0.06)	[−0.13, 0.11]	.849	0.03 (0.05)	[−0.06, 0.12]	.536
PCE accuracy	0.31 (0.36)	[−0.40, 1.01]	.397	−0.10 (0.28)	[−0.64, 0.45]	.728
Year (age group)						
5 vs. 1	−0.95 (0.29)	[−1.51, −0.39]	**.001**	−0.38 (0.19)	[−0.75, −0.00]	*.048*
5 vs. 1	−1.07 (0.31)	[−1.68, −0.46]	**.001**	−0.17 (0.18)	[−0.52, 0.17]	.328
5 vs. 3	−0.12 (0.27)	[−0.64, 0.41]	.665	0.21 (0.20)	[−0.20, 0.61]	.314
Familiarity	−0.00 (0.06)	[−0.12, 0.11]	.937	0.04 (0.05)	[−0.05, 0.13]	.388
PCE time spent together	0.01 (0.02)	[−0.03, 0.04]	.655	0.01 (0.01)	[−0.02, 0.03]	.547
Year (age group)						
5 vs. 1	−0.94 (0.28)	[−1.49, −0.38]	**.001**	−0.39 (0.19)	[−0.77, −0.02]	*.040*
5 vs. 1	−1.03 (0.30)	[−1.62, −0.44]	**.001**	−0.22 (0.18)	[−0.58, 0.13]	.213
5 vs. 3	−0.09 (0.27)	[−0.62, 0.43]	.731	0.17 (0.20)	[−0.23, 0.57]	.403
Familiarity	−0.00 (0.06)	[−0.12, 0.11]	.960	0.04 (0.05)	[−0.05, 0.13]	.378

Note. The *p* values that are significant after Bonferroni correction are highlighted in bold, and the *p* values that would be significant without Bonferroni correction are highlighted in italics. PCE = perceptual crossing experiment; ESM = experience sampling methodology; B = unstandardized coefficient; SE = standard error; CI = confidence interval.

Table 3. Relationship Between PCE Variables and Self-Reported Social Skills (VPV).

	VPV		
	B (*SE*)	95% CI	*p*
PCE awareness	1.75 (0.85)	[0.08, 3.42]	*.040*
Year (age group)			
2 vs. 1	−9.46 (3.46)	[−16.24, −2.68]	**.006**
5 vs. 1	−6.36 (3.11)	[−12.46, −0.26]	*.041*
5 vs. 3	3.10 (2.99)	[−2.76, 8.96]	.299
Familiarity	0.24 (0.68)	[−1.10, 1.57]	.729
PCE accuracy	−2.28 (3.75)	[−9.63, 5.08]	.544
Year (age group)			
2 vs. 1	−10.29 (3.53)	[−17.21, −3.37]	**.004**
5 vs. 1	−7.07 (3.19)	[−13.33, −0.81]	*.027*
5 vs. 3	3.22 (3.08)	[−2.81, 9.26]	.295
Familiarity	0.46 (0.69)	[−0.89, 1.81]	.504
PCE time spent together	0.13 (0.18)	[−0.23, 0.49]	.487
Year (age group)			
2 vs. 1	−10.03 (3.45)	[−16.80, −3.27]	**.004**
5 vs. 1	−7.67 (3.10)	[−13.75, −1.59]	**.013**
5 vs. 3	2.36 (3.03)	[−3.57, 8.30]	.435
Familiarity	0.37 (0.68)	[−0.97, 1.70]	.589

Note. The *p* values that are significant after Bonferroni correction are highlighted in bold, and the *p* values that would be significant without Bonferroni correction are highlighted in italics. PCE = perceptual crossing experiment; VPV = vragenlijst psychosociale vaardigheden; B = unstandardized coefficient; SE = standard error; CI = confidence interval.

association between both PCE accuracy and PCE time spent together and self-reported social skills (Table 3). PCE awareness was statistically significant and positively associated with self-reported skills, such that the higher the reported awareness of the PCE interaction, the higher the self-reported social skills. This effect fell short of significance after Bonferroni correction ($\alpha = .05/3 = .017$). The late- and mid-adolescent age groups also reported, on average, lower social skills compared with the youngest group.

H3: Association Between Self-Reported Social Skills (VPV) and Social Interaction in Daily Life (ESM). We found no statistically significant association between self-reported social skills and being in company in daily life. There was a significant association between age group and being in company, such that, on average, older adolescents were less likely to report being in company in daily life. There was a statistically significant, positive association between self-reported social skills and social experience in daily life, such that the higher the social skills score, the higher the reported social experiences in daily life (Table 4). Bonferroni correction was not applicable here.

H4: Comparison PCE Social Capacity and Self-Reported Social Skills (VPV). To compare PCE capacity for social contingency detection and self-reported social skills (VPV) in the association with social interaction in daily life, VPV score was added to each of the three separately tested associations of PCE variables (awareness, accuracy, time spent together) with daily life social interaction, resulting in six tested associations (Table 5). The association between self-reported social skills and social experience in daily life remained when adjusted for PCE variables, also after Bonferroni correction ($\alpha = .05/6 = .008$).

Exploratory Findings

Interaction With Age Group. A significant interaction effect was found for the late adolescent group in the association between PCE awareness and social experience in daily life, such that this group showed a stronger increase of social experience in daily life compared with the other age groups. Significant interaction effects were also found for each age group in the association between self-reported social skills and social experience in daily life. That is, the mid-adolescent age group showed a stronger increase of social experience in daily life compared with the other age groups. The results of all tested interactions are presented in Supplemental Material C.

Repeated Analyses Including Only Moments of Interaction With the Company in Daily Life. In 83.6% of the ESM prompts when participants indicated being in company, they also indicated they were interacting with this company, based on

Table 4. Relationship Between Self-Reported Social Skills (VPV) and Daily Life Social Interaction (ESM).

	ESM					
	Company			Social experience		
	B (*SE*)	95% CI	*p*	*B* (*SE*)	95% CI	*p*
VPV	0.01 (0.01)	[−0.00, 0.02]	.207	0.03 (0.01)	[0.02, 0.04]	<**.001**
Year (age group)						
3 vs. 1	−0.93 (0.25)	[−1.41, −0.44]	<**.001**	−0.03 (0.19)	[−0.41, 0.34]	.859
5 vs. 1	−1.04 (0.23)	[−1.49, −0.59]	<**.001**	0.15 (0.17)	[−0.19, 0.49]	.380
5 vs. 3	−0.11 (0.24)	[−0.57, 0.35]	.641	0.19 (0.19)	[−0.19, 0.56]	.329
Familiarity	−0.02 (0.06)	[−0.14, 0.09]	.693	0.02 (0.05)	[−0.07, 0.11]	.708

Note. The *p* values that are significant after Bonferroni correction are highlighted in bold. VPV = vragenlijst psychosociale vaardigheden; ESM = experience sampling methodology; B = unstandardized coefficient; SE = standard error; CI = confidence interval.

Table 5. Comparison Between PCE Capacity for Social Contingency Detection and Self-Reported Social Skills (VPV) in the Association With Social Interaction in Daily Life.

	ESM					
	Company			Social experience		
	B (*SE*)	95% CI	*p*	*B* (*SE*)	95% CI	*p*
PCE awareness	0.12 (0.08)	[−0.04, 0.29]	.137	0.10 (0.06)	[−0.02, 0.22]	.114
VPV	0.01 (0.01)	[−0.01, 0.02]	.366	0.03 (0.01)	[0.02, 0.04]	<**.001**
PCE accuracy	0.18 (0.35)	[−0.50, 0.87]	.600	−0.18 (0.27)	[−0.71, 0.36]	.514
VPV	0.01 (0.01)	[−0.00, 0.02]	.194	0.03 (0.01)	[0.02, 0.04]	<**.001**
PCE time spent together	0.00 (0.02)	[−0.03, 0.04]	.867	−0.00 (0.01)	[−0.03, 0.02]	.791
VPV	0.01 (0.01)	[−0.00, 0.02]	.212	0.03 (0.01)	[0.02, 0.04]	<**.001**

Note. The *p* values that are significant after Bonferroni correction are highlighted in bold. PCE = perceptual crossing experiment; VPV = vragenlijst psychosociale vaardigheden; ESM = experience sampling methodology; B = unstandardized coefficient; SE = standard error; CI = confidence interval.

a score of at least 2 on the item "We are interacting." When only including these moments in the analyses, overall, the confidence intervals around the regression coefficients became smaller. In particular, the differences between mid-adolescents and early adolescents became significant in the association with social experience in daily life. The results of these analyses are presented in Supplemental Material C.

Discussion

Within the scope of the current investigation, our findings did not provide evidence to support ecological and convergent validity for the PCE's measure of the capacity for social contingency detection. While the associations between PCE awareness and social experience in daily life, and between PCE awareness and self-reported social skills (VPV) were in the expected (i.e., positive) direction, these were nonsignificant. In addition, no statistically significant associations were found between both PCE accuracy and PCE time spent together and social interaction in daily life. Similarly, no associations were found between PCE accuracy and PCE time spent together and self-reported social skills. We found a positive association between self-reported social skills (VPV) and social experience in daily life. This robust association between self-reported social skills (VPV) and social experience in daily life did not change when adjusted for the variance explained by PCE awareness. We discuss the findings in more detail below, and consider their implications for further investigations of the PCE's validity.

The Absence of an Association Between PCE Social Capacity and Social Interaction in Daily Life

We expected better PCE performance—reflecting better capacity for social contingency detection—to relate to increased social behavior, as well as a better experience of company in the real world, across different social contexts. However, we did not find evidence to support our hypothesis,

which may indicate that what we measure in the PCE does not relate to daily life social interaction. Yet, previous studies demonstrated a heightened awareness during the PCE that was uniquely associated with reciprocity as opposed to non-reciprocity, in both adults and adolescents (e.g., Froese, Iizuka, & Ikegami, 2014; Froese et al., 2020; Hermans et al., 2020). We argue that the complete absence of a link between PCE performance and daily life social interaction is unlikely, as the PCE in these studies has been shown to capture reciprocity, which is fundamental to social interaction. We describe two potential methodological explanations for not finding an association between PCE performance and daily life social interaction.

First, while the translation of experimentally assessed capacity for social contingency detection did not translate to daily life social interaction, this may be a function of the way in which we measured social interaction, that is, the specific ESM items we used to operationalize "daily-life social interaction." As the PCE was primarily designed to study the relationship between sensorimotor interaction and social awareness (Auvray et al., 2009), finding no association between PCE performance and social behavior in daily life may not be entirely unexpected. In particular, adolescents may not always be able to choose when they want company or not (R. Achterhof, Kirtley, et al., 2020). Instead, PCE performance could be more directly associated with the level of subjective social *experience* in daily life, as opposed to the amount of social *behavior*. That is, adolescents who are more sensitive to the responsive presence of others, even under the minimal conditions of the PCE, could be expected to score higher on social skills, and rate social experience in daily life more highly. Indeed, although not statistically significant following Bonferroni correction, the direction of our results was consistent with this idea. An additional consideration is that the ESM social experience variable we used was averaged within and across individuals. A variable capturing within-person variability across moments may have been more appropriate to test associations with a dynamic adaptability score obtained by the PCE.

Second, our findings suggest that the PCE uniquely captures the capacity for social contingency detection within a specific PCE interaction, which only results from real-time interaction itself. This is essentially what dynamic systems theory would predict: that the simultaneous, responsive behavior of two individuals with the same goal results in a better coordination than what would be expected from the mere sum of isolated behavior (Froese & Gallagher, 2012; Froese, Iizuka, & Ikegami, 2013; Thelen & Smith, 2006). Indeed, recent simulation studies have demonstrated that coordination of movement (reciprocity) resulted in enhanced complexity that could be reached neither by studying independent, isolated behaviors nor in studying communication with a nonresponsive social stimulus (Candadai et al., 2019; Froese et al., 2020). Therefore, another plausible interpretation for the absence of an association between PCE performance and daily life social interaction is that the PCE performance only reflects the specific interaction measured within that specific dyad. Consequently, the capacity for social contingency detection as manifested within this specific interaction would not necessarily relate to the capacity for social contingency detection as manifested within different dyads in the real world. If this interpretation of the findings is true, it may be argued that the PCE in the current setup did not measure what we set out to measure, as we aimed to capture a context-independent capacity to flexibly adapt to changing social situations.

If each captured interaction is dependent on the specific context and dyad, the next question is how to establish ecological validity of the PCE. Averaging PCE performance within one individual in interaction with more than one other individual, that is, testing multiple dyads, could elucidate the extent to which the capacity for social contingency detection captured with the PCE is context-dependent, which will inform future use of the PCE. Examples of future setups include participants interacting with a computer with varying levels of adaptation to the participant, or participants switching dyads in a prolonged experiment. Both could result in an adaptability score within the PCE, which may be closer to what we ultimately aim for: predicting social adaptability to fluctuating social environments throughout the day, using an individual adaptability score. Therefore, we would expect this average adaptability score to be associated with daily life social interaction, or with individual variability of social behavior and experience in daily life.

Higher Self-Reported Social Skills Are Associated With a Better Social Experience in Daily Life

Self-reports of better social skills were robustly associated with better social experiences in daily life, although these were not associated with social behavior in daily life. Social behavior itself may be less affected by social skills than the experience of social interactions, at least in the general population. In particular, adolescents typically spend the majority of their time in structured situations such as school (R. Achterhof, Kirtley, et al., 2020), giving them less freedom to choose whether they want to be with others. This decreases with older age when adolescents become increasingly independent from their caregivers and peers (Lee et al., 2018), in addition to experiencing being alone as more positive than younger adolescents (Vanhalst et al., 2013; Wang et al., 2013). This would be in line with the negative effect of age we found on the average number of reports in company. The association we found between self-reported social skills and social experience in daily life

indicates that individuals with higher self-reported social skills experience social interaction more positively and potentially as more rewarding compared with individuals with lower self-reported social skills. This is an interesting avenue for future work, especially given that social reward and motivation deficits are characteristic of psychopathologies with a strong social component, such as autism spectrum disorder (Bottini, 2018) and psychosis (Frost & Strauss, 2016). For instance, early detection and prevention of social impairments could focus on improving social experience in daily life.

Notably, as we found variation in self-reported social skills to relate to differences in social experience of interactions in daily life, we could similarly expect self-reported social skills to relate to awareness of the interaction within the PCE. That is, awareness of the PCE interaction reflects the experience of social responsiveness. Indeed, we found an association between PCE awareness and social experience in daily life in the unadjusted analyses, but this was no longer significant following the application of the highly conservative Bonferroni correction (Streiner & Norman, 2011). The limited information provided by PCE performance based on the interaction with only one other individual may explain this. However, this needs further scrutiny, especially given that the novelty of our research (in adolescence) precludes us from comparing effect sizes with previous studies. Nevertheless, testing the interaction across multiple dyads could be a future avenue to investigate awareness of interaction across different social interactions. The use of dyadic analysis to maximize the information obtained from these dyads may even further advance future validation efforts of the PCE, and these have been implemented in a recent adult PCE study (Froese et al., 2020).

Limitations and Future Considerations

Our findings show that the PCE requires future research to further explore its experimental value in social interaction research. In particular, averaging PCE scores across multiple dyads may be a starting point for future research. This recommendation could benefit from a laboratory environment (as opposed to the current school setting), in which other factors potentially impacting PCE performance, such as noise, can be minimized. Taking a measure of the capacity for social contingency detection back to the lab seems counter-intuitive, given our aim to improve the ecological validity of these measures. However, laboratory testing of real-time social interaction within dyads may focus on what is fundamentally underlying social interaction, while still addressing the dynamic and continuous coordination among individuals. More controlled testing would also allow further refinement of the PCE variables.

Furthermore, the PCE is assessing dynamic interactions between two individuals. Still, we use this task to draw inferences about the *individual* capacity for social contingency detection. However, the task performance is obviously influenced by the capacity of the other individual as well. Interacting with someone who is performing more poorly may influence the overall dynamic interaction in a negative way, while interacting with someone who is particularly good at the task, may improve the overall dynamic interaction substantially. One way of improving this would be to allow the individuals to participate in multiple dyads and calculate the average performance of the individual over these different dyads.

Another step toward validation of the PCE includes testing associations with social skills as assessed by external raters, such as parents or teachers. This would also address the potential impact of response styles, such that the correlation between different self-report measures is related to common method effects (Paulhus, 1991). Another important future validation step is testing associations between PCE performance and other tasks that aim to measure real-time social interaction. In particular, tasks with a focus on another modality than touch, such as the visual modality in studies on dyadic gaze behavior (Hessels, 2020), are interesting in this respect. Investigating other modalities than verbal interaction in social interaction research may even elucidate ways to capture interaction below the level of explicit awareness.

Our study operationalized "social experience" as a sum score of the variables, "I feel at ease in this company," "I feel appreciated in this company," and "I belong (in this company)." However, this does not exhaustively capture all aspects of daily life social experience. Future research may benefit from investigating other relevant aspects of social experience, for example, social acceptance or (friendship) closeness (Pouwels et al., 2021).

In addition, ESM compliance varied substantially, and unlike previous ESM studies in adults (e.g., Palmier-Claus et al., 2011), we did not exclude participants if their compliance rate was lower than one third of beeps to avoid potentially biasing our inferences (Jacobson, 2020). Although our sample's average compliance rate was higher than the average compliance in the entire SIGMA study sample (51% vs. 41%, based on participants who completed at least one beep), it still reflects the lower end of the range of compliance rates reported by other ESM studies in adolescents (van Roekel et al., 2019). Our adolescent population spent a significant amount of time at school (43% in the entire sample; Kirtley et al., 2019), which is also the moment that notifications are generally experienced as least inconvenient by adolescents (van Roekel et al., 2019). However, in contrast to findings by Van Roekel et al. (2019), we found that late adolescents completed significantly fewer beeps than the younger age groups. It may be the case that adolescents completed fewer beeps when they were in the company of others compared with when they were alone. This would

have resulted in limited variability in the social behavior and experience variables, which could impact the strength of associations and introduce bias. Therefore, this limitation should be considered in the interpretation of results and, if possible, prevented in future research.

Finally, the absence of a separate power analysis for the current study's sample—due to no available previous parameter estimates—can be regarded as a limitation and should be considered when interpreting the findings. Future PCE studies could use the current study's estimates to conduct *a priori* power analyses. Similarly, future PCE studies should attempt to replicate the psychometric properties of the novel scales used in the current study. Following Flake et al. (2017), the use of Cronbach's alpha to test the reliability of items, and the use of short scales to capture our constructs may be considered methodologically suboptimal. However, our current measures were grounded in previous PCE research and the computation of our variables has been preregistered. Therefore, we argue that the current study's purpose to evaluate the PCE's validity with regard to external measures, including previously used short scales, is valid at this stage.

Conclusion

The current study is the first in which the capacity for social contingency detection, as captured with the PCE, has been studied in relation to other external measures of social skills and behaviors. No significant associations were found between PCE performance and (a) the probability and experience of being in company during daily life, and (b) self-reported social skills. Within the scope of these specific associations, therefore, we did not find evidence supporting ecological and convergent validity of the PCE. Although our results were in the expected direction, we only found a positive association between self-reported social skills and social experience of interactions in daily life. Yet, our findings have also uncovered novel hypotheses for future work in social interaction research, particularly regarding testing multiple dyads to obtain a social adaptability score that may be predictive of daily life social interaction.

Given the importance of early detection of social impairments and the quickly changing social environments to which adolescents must adapt (Dahl et al., 2018), we continue to encourage future work to include social dynamics as opposed to individualistic methods of the capacity for social contingency detection. The PCE is the first experimental task used in adolescents in which the crucial interactive element to simulate real social interaction has been addressed from an interactionist point of view (Schilbach, 2016). As the complexity of capturing social dynamics should not be underestimated (Fried & Robinaugh, 2020), we argue that future work to improve the PCE setup and investigate the validity of the experiment is worthwhile.

Author Note

The experimental data used for the current manuscript have previously been used for two other studies (Hermans et al., 2020, 2021). Descriptive data on the SIGMA project, where the current study's sample is part of, are published in a (Dutch) national report (Kirtley et al., 2019), available via www.sigma-leuven.be. This manuscript has been published as a preprint on PsyArXiv, available via https://psyarxiv.com/jrzbw/.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data Availability Statement

The data sets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request. The data are not publicly available due to privacy or ethical restrictions. Experience Sampling Methodology (ESM) items are publicly available in the ESM Item Repository (Kirtley et al., 2019). The study was preregistered at the website of the Open Science Framework, available via https://osf.io/bxzpk. Discrepancies between the preregistration and the final report are detailed in a transparent changes document. This document and the STATA code used for the analyses are available via the OSF website: https://osf.io/nhs6d/?view_only=4f7ea37b3cf648bdb7da48b1ce29dbb5. This study has been submitted as a preprint on PsyArXiv, available via https://psyarxiv.com/jrzbw/.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the FWO Odysseus grant to I.M.-G. (G0F8416N), which includes the PhD studentship of K.S.F.M.H. O.J.K. is supported by a senior postdoctoral fellowship from Research Foundation Flanders (FWO: 1257821N). Z.K. was supported by the European Union's Horizon 2020 research and innovation program under grant agreement no. 777084. R.F. and L.Z.-F. acknowledge financial support from the Universidad Nacional Autónoma de México (UNAM) through the DGAPA-PAPIIT project IA102619. L.Z.-F. is grateful for the scholarship 638215 granted by the CONACyT and a scholarship in the program Research Grants-Bi-nationally Supervised Doctoral Degrees, 2019/20 granted by the German Academic Exchange Service (DAAD, Deutscher Akademischer Austauschdienst).

Methodological Disclosure

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study.

ORCID iD

Karlijn S. F. M. Hermans https://orcid.org/0000-0003-3570-6479

Supplemental Material

Supplemental material for this article is available online.

Notes

1. Although these components resemble the sensory, cognitive-interpretative, and motoric aspects of reciprocal behavior in the terminology, we adopted from Constantino et al. (2000), and we emphasize the dynamic interplay of these components during interaction (in line with dynamic systems theory), instead of the isolated expression of each component in a certain order (in line with traditional cognitivist theories), as also set out by Froese and Gallagher (2012).
2. The education system in Flanders allows for adolescents to redo a year in school. With age groups based on school year, this results in heterogeneous age groups, which show some overlap (Table 1).
3. The term "social experience" has been adopted from previous Experience Sampling Methodology (ESM) papers, distinguishing between social experiences and social behavior (Achterhof, Kirtley, et al., 2021, May 21; Achterhof, Kirtley, et al., 2020; Achterhof, Kirtley, et al., 2021; Achterhof, Myin-Germeys, et al., 2021, June 14; Achterhof, Kirtley, et al., 2020). Social experience captures a broad range of variables tapping into different experiential aspects of the social encounter, including social comfort and social acceptance.
4. gamma_{0ij} represents the mean of individual *i* on school *j*, and the error e_tij represents the deviation from the participant's outcome on school *j* at time *t*. The participant's mean level as denoted by b00i and 0ji is the deviation of the day mean level for this participant from their mean level. b000 denotes the grand mean of the outcome for the population, and u_j represents the deviation of the mean level for the participants in school *j* from the population mean. In this model, the Level 1 errors are normally distributed. The random effects v_{ij} and u_j are normally distributed.
5. In error, this hypothesis was not preregistered. As this analysis was a prerequisite for testing H4, we have still included it here in the primary analysis section. Although, we recognize that, its non-preregistration means it is not strictly confirmatory.
6. In Supplemental Material B, the results are presented including 18 participants with 1, 2, or 3 missing items on the VPV for whom data were imputed (using the average person VPV score), in addition to a sensitivity analysis.

References

Achterhof, R., Kirtley, O. J., Schneider, M., Hagemann, N., Hermans, K. S. F. M., Hiekkaranta, A. P., Lecei, A., Decoster, J., Derom, C., De Hert, M., Guloksuz, S., Jacobs, N., Menne-Lothmann, C., Rutten, B., Thiery, E., van Os, J., van Winkel, R., Wichers, M., & Myin-Germeys, I. (2020, August 11). General psychopathology and its social correlates in the daily lives of youth. *PsyArXiv*. Advance online publication.

Achterhof, R., Kirtley, O. J., Schneider, M., Hagemann, N., Hermans, K. S. F. M., Hiekkaranta, A. P., Lecei, A., Lafit, G., & Myin-Germeys, I. (2022). Adolescents' real-time social and affective experiences of online and face-to-face interactions. *Computers in Human Behavior*, 129, Article 107159. https://doi.org/10.1016/j.chb.2021.107159

Achterhof, R., Kirtley, O. J., Schneider, M., Lafit, G., Hagemann, N., Hermans, K. S. F. M., Hiekkaranta, A. P., Lecei, A., & Myin-Germeys, I. (2021). Daily-life social experiences as a potential mediator of the relationship between parenting and psychopathology in adolescence. *Frontiers in Psychiatry*, *12*, Article 697127. https://doi.org/10.3389/fpsyt.2021.697127

Achterhof, R., Myin-Germeys, I., Bamps, E., Hagemann, N., Hermans, K. S. F. M., Hiekkaranta, A. P., Janssens, J., Lecei, A., Lafit, G., & Kirtley, O. (2021, June 14). COVID-19-related changes in adolescents' daily-life social interactions and psychopathology symptoms. *PsyArXiv*. Advance update publication. https://doi.org/10.31234/osf.io/5nfp2

Achterhof, R., Schneider, M., Kirtley, O. J., Wampers, M., Decoster, J., Derom, C., & Myin-Germeys, I. (2020, March 23). Be (com)ing social: Daily-life social interactions and parental bonding. *PsyArXiv*. Advance online publication. https://doi.org/10.31234/osf.io/xwkdq

Anderson, S. F., Kelley, K., & Maxwell, S. E. (2017). Sample-size planning for more accurate statistical power: A method adjusting sample effect sizes for publication bias and uncertainty. *Psychological Science*, *28*(11), 1547–1562. https://doi.org/10.1177/0956797617723724

Auvray, M., Lenay, C., & Stewart, J. (2009). Perceptual interactions in a minimalist virtual environment. *New Ideas in Psychology*, *27*(1), 32–47. https://doi.org/10.1016/j.newideapsych.2007.12.002

Auvray, M., & Rohde, M. (2012). Perceptual crossing: The simplest online paradigm. *Frontiers in Human Neuroscience*, *6*, 181. https://doi.org/10.3389/fnhum.2012.00181

Benning, S. D., Bachrach, R. L., Smith, E. A., Freeman, A. J., & Wright, A. G. C. (2019). The registration continuum in clinical science: A guide toward transparent practices. *Journal of Abnormal Psychology*, *128*(6), 528–540. https://doi.org/10.1037/abn0000451

Birchwood, M., Smith, J., Cochrane, R., Wetton, S., & Copestake, S. (1990). The social functioning scale the development and validation of a new scale of social adjustment for use in family intervention programmes with schizophrenic patients. *British Journal of Psychiatry*, *157*(6), 853–859. https://doi.org/10.1192/bjp.157.6.853

Blakemore, S.-J., & Mills, K. L. (2014). Is adolescence a sensitive period for sociocultural processing? *Annual Review of Psychology*, *65*(1), 187–207. https://doi.org/10.1146/annurev-psych-010213-115202

Bottini, S. (2018). Social reward processing in individuals with autism spectrum disorder: A systematic review of the social motivation hypothesis. *Research in Autism Spectrum Disorders*, *45*, 9–26. https://doi.org/10.1016/j.rasd.2017.10.001

Candadai, M., Setzler, M., Izquierdo, E. J., & Froese, T. (2019). Embodied dyadic interaction increases complexity of neural dynamics: A minimal agent-based simulation model. *Frontiers in Psychology*, *10*, 540. https://doi.org/10.3389/fpsyg.2019.00540

Constantino, J. N., Przybeck, T., Friesen, D., & Todd, R. D. (2000). Reciprocal social behavior in children with and without pervasive developmental disorders. *Journal of Developmental & Behavioral Pediatrics*, *21*(1), 2–11.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302. https://doi.org/10.1037/h0040957

Csikszentmihalyi, M., & Larson, R. (2014). Validity and reliability of the experience-sampling method. In: *Flow and the foundations of positive psychology* (pp. 35–54). Springer, Dordrecht.. https://doi.org/10.1007/978-94-017-9088-8_3

Dahl, R. E., Allen, N. B., Wilbrecht, L., & Suleiman, A. B. (2018). Importance of investing in adolescence from a developmental science perspective. *Nature*, *554*(7693), 441–450. https://doi.org/10.1038/nature25770

De Jaegher, H., Di Paolo, E., & Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Sciences*, *14*(10), 441–447. https://doi.org/10.1016/j.tics.2010.06.009

Del Prette, Z. A. P., Prette, A. D., De Oliveira, L. A., Gresham, F. M., & Vance, M. J. (2012). Role of social performance in predicting learning problems: Prediction of risk using logistic regression analysis. *School Psychology International*, *33*(6), 615–630. https://doi.org/10.1177/0020715211430373

Flake, J. K., Pek, J., & Hehman, E. (2017). Construct validation in social and personality research. *Social Psychological and Personality Science*, *8*(4), 370–378. https://doi.org/10.1177/1948550617693063

Fried, E. I., & Robinaugh, D. J. (2020). Systems all the way down: Embracing complexity in mental health research. *BMC Medicine*, *18*(1), Article 205. https://doi.org/10.1186/s12916-020-01668-w

Froese, T., & Gallagher, S. (2012). Getting interaction theory (IT) together: Integrating developmental, phenomenological, enactive, and dynamical approaches to social interaction. *Interaction Studies*, *13*(3), 436–468. https://doi.org/10.1075/is.13.3.06fro

Froese, T., Iizuka, H., & Ikegami, T. (2013). From synthetic modeling of social interaction to dynamic theories of brain–body–environment–body–brain systems. *Behavioral and Brain Sciences*, *36*(4), 420–421. https://doi.org/10.1017/s0140525x12001902

Froese, T., Iizuka, H., & Ikegami, T. (2014). Embodied social interaction constitutes social cognition in pairs of humans: A minimalist virtual reality experiment. *Science Reports*, *4*(1), Article 3672. https://doi.org/10.1038/srep03672

Froese, T., Zapata-Fonseca, L., Leenen, I., & Fossion, R. (2020). The feeling is mutual: Clarity of haptics-mediated social perception is not associated with the recognition of the other, only with recognition of each other. *Frontiers in Human Neuroscience*, *14*, 560567. https://doi.org/10.3389/fnhum.2020.560567

Frost, K. H., & Strauss, G. P. (2016). A review of anticipatory pleasure in schizophrenia. *Current Behavioral Neuroscience Reports*, *3*, 232–247. https://doi.org/10.1007/s40473-016-0082-5

Gresham, F. M. (2016). Social skills assessment and intervention for children and youth. *Cambridge Journal of Education*, *46*(3), 319–332. https://doi.org/10.1080/0305764x.2016.1195788

Hermans, K. S. F. M., Kasanova, Z., Zapata-Fonseca, L., Lafit, G., Fossion, R., Froese, T., & Myin-Germeys, I. (2020). Investigating real-time social interaction in pairs of adolescents with the Perceptual Crossing Experiment. *Behavior Research Methods*, *52*(5), 1929–1938. https://doi.org/10.3758/s13428-020-01378-4

Hermans, K. S. F. M., Kirtley, O. J., Kasanova, Z., Achterhof, R., Hagemann, N., Hiekkaranta, A. P., Lecei, A., Zapata-Fonseca, L., Lafit, G., Fossion, R., Froese, T., & Myin-Germeys, I. (2021). Capacity for social contingency detection continues to develop across adolescence. *Social Development*, 1–19. https://doi.org/10.1111/sode.12567

Hessels, R. S. (2020). How does gaze to faces support face-to-face interaction? A review and perspective. *Psychonomic Bulletin & Review*, *27*, 856–881. https://doi.org/10.3758/s13423-020-01715-w

Hox, J. J., Moerbeek, M., & Van de Schoot, R. (2010). *Multilevel analysis: Techniques and applications*. Routledge.

Jacobson, N. C. (2020). *Compliance thresholds in intensive longitudinal data: Worse than listwise deletion: Call for action* [Symposium]. Society for Ambulatory Assessment, Melbourne, Australia.

Kessler, R. C. (2007). Psychiatric epidemiology: Challenges and opportunities. *International Review of Psychiatry*, *19*(5), 509–521. https://doi.org/10.1080/09540260701564914

Kirtley, O. J., Achterhof, R., Hagemann, N., Hermans, K. S. F. M., Hiekkaranta, A. P., Lecei, A., Boets, B., Henquet, C., Kasanova, Z., Schneider, M., Van Winkel, R., Reininghaus, U., Viechtbauer, W., & Myin-Germeys, I. (2021, April 2). Initial cohort characteristics and protocol for SIGMA: An accelerated longitudinal study of environmental factors, inter- and intrapersonal processes, and mental health in adolescence. *PsyArXiv*. Advance online publication. https://doi.org/10.31234/osf.io/jp2fk

Kirtley, O. J., Hiekkaranta, A. P., Kunkels, Y. K., Verhoeven, D., Van Nierop, M., & Myin-Germeys, I. (2019). *The Experience Sampling Method (ESM) item repository*. https://doi.org/10.17605/OSF.IO/KG376

Lafit, G., Adolf, J. K., Dejonckheere, E., Myin-Germeys, I., Viechtbauer, W., & Ceulemans, E. (2021). Selection of the number of participants in intensive longitudinal studies: A user-friendly Shiny app and tutorial to perform power analysis in multilevel regression models that account for temporal dependencies. *Advances in Methods and Practices in Psychological Science*, *4*(1), Article 2515245920978738. https://doi.org/10.1177/2515245920978738

Laurenceau, J.-P., & Bolger, N. (2012). Analyzing diary and intensive longitudinal data from dyads. In M. Mehl & T. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 407–422). Guilford.

Lee, K. T. H., Lewis, R. W., Kataoka, S., Schenke, K., & Vandell, D. L. (2018). Out-of-school time and behaviors during adolescence. *Journal of Research on Adolescence*, *28*(2), 284–293. https://doi.org/10.1111/jora.12389

Lenay, C., & Stewart, J. (2012). Minimalist approach to perceptual interactions. *Frontiers in Human Neuroscience*, *6*, 98. https://doi.org/10.3389/fnhum.2012.00098

Leong, V., & Schilbach, L. (2019). The promise of two-person neuroscience for developmental psychiatry: Using interaction-based sociometrics to identify disorders of social interaction. *British Journal of Psychiatry*, *215*(5), 636–638. https://doi.org/10.1192/bjp.2019.73

Little, S. G., Swangler, J., & Akin-Little, A. (2017). Defining social skills. In J. Matson (Ed.), *Handbook of social behavior and skills in children* (pp. 9–17). Springer. https://doi.org/10.1007/978-3-319-64592-6_2

Meers, K., Dejonckheere, E., Kalokerinos, E. K., Rummens, K., & Kuppens, P. (2020). MobileQ: A free user-friendly application for collecting experience sampling data. *Behavior Research Methods*, *52*, 1510–1515. https://doi.org/10.3758/s13428-019-01330-1

Mote, J., & Fulford, D. (2020). Ecological momentary assessment of everyday social experiences of people with schizophrenia: A systematic review. *Schizophrenia Research*, *216*, 56–68. https://doi.org/10.1016/j.schres.2019.10.021

Myin-Germeys, I. (2020). Real-life social interactions in schizophrenia. *Schizophrenia Research*, *216*, 10–11. https://doi.org/10.1016/j.schres.2019.11.042

Myin-Germeys, I., Kasanova, Z., Vaessen, T., Vachon, H., Kirtley, O., Viechtbauer, W., & Reininghaus, U. (2018). Experience sampling methodology in mental health research: New insights and technical developments. *World Psychiatry*, *17*(2), 123–132.

Newen, A., De Bruin, L., & Gallagher, S. (Eds.). (2018). *The Oxford handbook of 4E cognition*. Oxford University Press.

Osborne-Crowley, K. (2020). Social cognition in the real world: Reconnecting the study of social cognition with social reality. *Review of General Psychology*, *24*(2), 144–158. https://doi.org/10.1177/1089268020906483

Palmier-Claus, J. E., Myin-Germeys, I., Barkus, E., Bentley, L., Udachina, A., Delespaul, P. A. E. G., Lewis, S. W., & Dunn, G. (2011). Experience sampling research in individuals with mental illness: Reflections and guidance. *Acta Psychiatrica Scandinavica*, *123*(1), 12–20. https://doi.org/10.1111/j.1600-0447.2010.01596.x

Pan, X., & Hamilton, A. F. C. (2018). Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape. *British Journal of Psychology*, *109*(3), 395–417. https://doi.org/10.1111/bjop.12290

Paulhus, D. L. (1991). *Measurement and control of response bias*. Academic Press. https://doi.org/10.1016/B978-0-12-590241-0.50006-X

Pouwels, J. L., Valkenburg, P. M., Beyens, I., van Driel, I. I., & Keijsers, L. (2021). Social media use and friendship closeness in adolescents' daily lives: An experience sampling study. *Developmental Psychology*, *57*(2), 309–323.

Rose, L. T., & Fischer, K. W. (2008). Dynamic systems theory. In R. A. Shweder (Ed.), *Chicago companion to the child*. University of Chicago Press. https://www.gse.harvard.edu/~ddl/articlesCopy/RoseFischer_DynamicSystems.ChicagoCompanion2008.pdf.

Schilbach, L. (2016). Towards a second-person neuropsychiatry. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150081. https://doi.org/10.1098/rstb.2015.0081

Schneider, M., Reininghaus, U., Van Nierop, M., Janssens, M., & Myin-Germeys, I. (2017). Does the Social Functioning Scale reflect real-life social functioning? An experience sampling study in patients with a non-affective psychotic disorder and healthy control individuals. *Psychological Medicine*, *47*(16), 2777–2786. https://doi.org/10.1017/s0033291717001295

StataCorp. (2015). *Stata statistical software: Release 14*. StataCorp LP.

Streiner, D. L., & Norman, G. R. (2011). Correction for multiple testing. *Chest*, *140*(1), 16–18. https://doi.org/10.1378/chest.11-0523

Thelen, E., & Smith, L. B. (2006). Dynamic systems theories. In W. Damon & R. M. Lerner (Series Ed.), *Handbook of child psychology* (6th ed., 258–312). https://doi.org/10.1002/9780470147658.chpsy0106

Trevarthen, C., & Aitken, K. J. (2001). Infant intersubjectivity: Research, theory, and clinical applications. *Journal of Child Psychology and Psychiatry*, *42*(1), 3–48. https://doi.org/10.1111/1469-7610.00701

Van der Ploeg, J. D., & Scholte, E. M. (2013). *Vragenlijst Psychosociale Vaardigheden*. Bohn Stafleu van Loghum.

Vanhalst, J., Goossens, L., Luyckx, K., Scholte, R. H. J., & Engels, R. C. M. E. (2013). The development of loneliness from mid- to late adolescence: Trajectory classes, personality traits, and psychosocial functioning. *Journal of Adolescence*, *36*(6), 1305–1312. https://doi.org/10.1016/j.adolescence.2012.04.002

van Roekel, E., Keijsers, L., & Chung, J. M. (2019). A review of current ambulatory assessment studies in adolescent samples and practical recommendations. *Journal of Adolescent Research*, *29*(3), 560–577. https://doi.org/10.1111/jora.12471

Wang, J. M., Rubin, K. H., Laursen, B., Booth-Laforce, C., & Rose-Krasnor, L. (2013). Preference-for-solitude and adjustment difficulties in early and late adolescence. *Journal of Clinical Child & Adolescent Psychology*, *42*(6), 834–842. https://doi.org/10.1080/15374416.2013.794700

Zampella, C. J., Csumitta, K. D., Simon, E., & Bennetto, L. (2020). Interactional synchrony and its association with social and communication ability in children with and without autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *50*, 3195–3206. https://doi.org/10.1007/s10803-020-04412-8

Zapata-Fonseca, L., Dotov, D., Fossion, R., & Froese, T. (2016). Time-series analysis of embodied interaction: Movement variability and complexity matching as dyadic properties. *Frontiers in Psychology*, *7*, 1940. https://doi.org/10.3389/fpsyg.2016.01940