Title: Investigating real-time social interaction in pairs of adolescents with the Perceptual Crossing Experiment

Authors: Karlijn S.F.M. Hermans[1], Zuzana Kasanova[1], Leonardo Zapata-Fonseca[2,3,4], Ginette Lafit[1,5], Ruben Fossion[3], Tom Froese[3,6], Inez Myin-Germeys[1]

[1]KU Leuven, Department of Neuroscience, Center for Contextual Psychiatry, Leuven, Belgium

[2]Plan of Combined Studies in Medicine (PECEM), Faculty of Medicine, National Autonomous University of Mexico, Mexico City, Mexico

[3]Center for the Sciences of Complexity (C3), National Autonomous University of Mexico, Mexico city, Mexico

[4]Section Phenomenological Psychopathology and Psychotherapy, Department of General Psychiatry, Center of Psychosocial Medicine, University of Heidelberg, Heidelberg, Germany

[5]KU Leuven, Department of Psychology, Research Group of Quantitative Psychology and Individual Differences, Leuven, Belgium

[6]Department of Computer Science, Institute of Applied Mathematics and Systems Research, National Autonomous University of Mexico, Mexico City, Mexico

Corresponding author: Karlijn Hermans, Center for Contextual Psychiatry, Department of Neuroscience, Catholic University of Leuven, Kapucijnenvoer 33, blok h, 3000 Leuven, Belgium; tel: +32-(0)16 3242 22; e-mail: karlijn.hermans@kuleuven.be

Word count abstract: 170

Word count main text: 5945 (including Figure legends)

Abstract

The study of real-time social interaction provides ecologically-valid insight into social behavior. The objective of the current research is to experimentally assess real-time social contingency detection in an adolescent population, using a shortened version of the Perceptual Crossing Experiment (PCE). Pairs of 148 adolescents aged between 12 and 19 were instructed to find each other in a virtual environment interspersed with other objects by interacting with each other using tactile feedback only. Across six rounds, participants demonstrated increasing accuracy in social contingency detection, which was associated with increasing subjective experience of the mutual interaction. Subjective experience was highest in rounds when both participants were simultaneously accurate in detecting each other's presence. The six-round version yielded comparable social contingency detection outcome measures to a ten-round version of the task. The shortened six-round version of the PCE has therefore enabled us to extend the previous findings on social contingency detection in adults to an adolescent population, allowing implementation in prospective research designs to assess development of social contingency detection over time.

Key words social interaction – ecological validity – social contingency detection – virtual paradigm - adolescence

1. Introduction

Research on the dynamics of social interaction and its assessment have been dominated by studies with a key role for social cognition, focusing on cognitive processes within one individual (Blakemore & Choudhury, 2006; Brizio, Gabbatore, Tirassa, & Bosco, 2015; Hutto, Herschbach, & Southgate, 2011). These entail for example the ability to understand others' mental states, or 'mentalizing', studied in both normal development (Bosco, Gabbatore, & Tirassa, 2014) and psychopathology (Penn, Sanna, & Roberts, 2008). However, findings from laboratory research on social cognition have only partly been able to explain social functioning (Fett et al., 2011; Simons, Bartels-Velthuis, & Pijnenborg, 2016), which questions the assumption that social cognition is a prerequisite for social functioning (Hermans et al., 2019; Schneider, Myin, & Myin-Germeys, 2019). Although many other factors can be assessed to increase the explained variance in social functioning (Barch, Pagliaccio, & Luking, 2016), it is argued that crucial information continues to be lacking if these factors only relate to one individual, i.e. the observer. As an alternative to the observer's point of view, the interactor's point of view reflects the ongoing interaction and constantly changing dynamics in the environment (Schilbach et al., 2013). Therefore, paradigms studying social interaction, using the mutual engagement of at least two parties, have been gaining increased attention (De Jaegher & Di Paolo, 2008; Froese, 2018). In the assessment of real-time interaction, both interactors engage at the same time, allowing to capture the dynamics of social interaction itself. This goes beyond the study of social cognition as internal mechanism, and instead focuses on basic capacities that constitute the dynamics of constantly changing behavior in interaction with body and environment.

From a very early age, infants show engagement with others and responsiveness to social cues (Reddy, 2010). The development of real-time responding to social cues and maintaining a social interaction has been studied in two-month-old babies, showing that infants could distinguish live from prerecorded interactions with their mother (Murray & Trevarthen, 1985; Nadel, Carchon, Kervella, Marcelli, & Reserbat-Plantey, 1999). The study of real-time successful interactive capacities has been continued in adults with the development of the Perceptual Crossing Experiment (PCE) (Auvray, Lenay, & Stewart, 2009; Auvray & Rohde, 2012). Within this experimental paradigm, interaction is defined as the co-regulated coupling between two autonomous individuals within their environment (De Jaegher, Di Paolo, & Gallagher, 2010). This is captured by the assessment of social contingency detection, which is defined as the sensitivity to other people's responsiveness to one's presence and behavior. Previous research has investigated social contingency detection in terms of objective measures of the interaction process, such as detection accuracy and turn-taking, and in terms of the participant's subjective experience of interaction (Froese, Iizuka, & Ikegami, 2014; Zapata-Fonseca, Dotov, Fossion, & Froese, 2016). Detection accuracy entailed correct detection of the other based on real-time interaction with tactile feedback only. Turn-taking reflected a strategy that participants employed in order to detect the other. Subjective experience of interaction was measured with self-report, assessing the experience of the interaction without being given feedback on detection accuracy.

The PCE has, to date, only been implemented in adults. These studies showed evidence for accurate mutual awareness through sensorimotor coordination in a minimal virtual environment (Auvray & Rohde, 2012; Deschamps, Lenay, Rovira, Le Bihan, & Aubert, 2016; Froese et al., 2014; Kojima, Froese, Oka, Iizuka, & Ikegami, 2017; Zapata-Fonseca et al., 2016). Positive associations have been found between the mutual interaction process and subjective experience thereof, suggesting that co-regulation of interaction in real time is necessary for successful detection of social contingency (Froese et al., 2014). Some authors have even argued on the basis of the PCE that social interaction may constitute social cognition (De Jaegher et al., 2010; Froese & Di Paolo, 2011). More recently, the PCE has also shown to distinguish interaction patterns, measured as amount and variability of movement towards each other, in individuals with high-functioning autism from those in controls (Zapata-Fonseca et al., 2019; Zapata-Fonseca, Froese, Schilbach, Vogeley, & Timmermans, 2018). These studies, showing the ability to capture variability in capacities for social contingency detection in populations with impairments in social functioning, provided further evidence for the capacity of the PCE to investigate mechanisms of social interaction.

Although the development of social capacities starts from birth, the role of interpersonal functioning is very important in adolescence, during which maturation of social skills takes place (Smetana, 2010; Smetana, Campione-Barr, & Metzger, 2006). This is also the age period in which disorders marked with changes in social behavior have their onset, such as depression, anxiety, and psychosis (Kelleher et al., 2012; Schilbach, 2016). Early detection of variability in social interaction styles associated with social impairments and their development could enable timely prevention and intervention efforts. In adolescent populations and prospective longitudinal designs, the rather extensive PCE previously used in adults - fifteen rounds following a training phase, lasting up to one hour (e.g. Froese et al., 2014; Zapata-Fonseca et al., 2018) - may benefit from a considerably shortened design. We therefore adapted and shortened the PCE to six rounds, and compared it with the original fifteen-round design with regard to the ability to assess the learning of social contingency detection. In addition, we conducted a ten-round version of the PCE in a subsample of adolescents in order to test whether social contingency detection improves or asymptotes after six rounds.

1.1 Research questions and hypotheses

The aim of this study is to assess the capacity for social contingency detection using a modified, shortened version of the Perceptual Crossing Experiment (PCE) in adolescents.

Accordingly, the main research questions are: Does the modified six-round version of the PCE assess 1) overall social contingency detection measured as amount of time spent together, correct detection of the other, and subjective experience of interaction in all rounds, and 2) learning of social contingency detection measured as an increase of levels of social contingency detection across six rounds; and 3) in a ten-round version of the PCE, does the average level of social contingency detection change in round seven to ten compared to round one to six? Specific hypotheses are detailed in supplementary material A.

2. Methods

2.1 Sample

Data collection took place between February 2018 and May 2019. Participants were recruited from the general population in Flanders, Belgium. This was done in secondary schools that participated in a large longitudinal cohort study on adolescent mental health and development; the SIGMA project (Kirtley et al., in preparation). Participation in the SIGMA project included 100 minutes of completing questionnaires in groups of 20 to 24 adolescents, from which eight were randomly selected to perform the PCE. The total number of participants was 148, of whom 116 completed six rounds and 32 completed ten rounds of the PCE (see procedure). Participants received a 10 Euro voucher after full participation in the SIGMA study. The exclusion criterion was an inadequate level of Dutch or English and therefore failure to understand the instructions. Ethical approval was provided for the entire SIGMA study protocol including the PCE (S6 1395). This study was registered prior to conducting the analyses, after the data collection was finished, on the website of the Open Science Framework (https://osf.io/jmbdr/?view_only=9206a27ca3834a7da8da116b6154d1ad). The preregistration adheres to the disclosure requirements of this institutional registry.

2.2 Experimental setup

Participants played a game together with a randomly assigned partner. They were instructed to imagine walking through a long, dark loop corridor with this partner. This is a virtual space not visible to the participants, in which they could move an avatar on one axis back and forth with their dominant hand, using a trackball. Their task was to find their partner's (i.e. the other) avatar in this space without communicating in any other way outside the interface. The setup, in which participants sat back to back and listened to Brownian noise via a headphone, prevented interactions outside the virtual space. Every time the participants' avatar overlapped with another entity in the virtual space, they received tactile feedback (i.e. a vibration) on the hand moving the trackball. Within the virtual space, participants could not only encounter the other avatar, but also a 'chair' and one 'other moving entity', which would each give exactly the same tactile feedback during an encounter. The other avatar is an animate, reactive entity. The 'chair' is an inanimate, non-reactive entity to be referred to as static object. The 'other moving entity' is an animate, non-reactive entity moving exactly as the other avatar but at a fixed distance of 150 pixels, to be referred to as the shadow. Each avatar could sense only its own static object, not the other's static object, which was positioned in a different location in virtual space (Figure 1).

The experimental setup was based on a previous study using the PCE in adults (Froese et al., 2014). The shared virtual space with connected endpoints consisted of 600 pixels and the entities (static object, shadow and other avatar) were 4 pixels wide (Figure 1). On the virtual line, the two static objects (one for each avatar) were fixed at 150 and 450 pixels. The distance between the avatar and the shadow was 150 pixels. The tactile feedback indicating an encounter involved vibration with a fixed intensity during the crossing of another entity, making the duration of the vibration dependent on how much time the avatars' pixels overlapped with this entity. Without overlap the vibration was off.

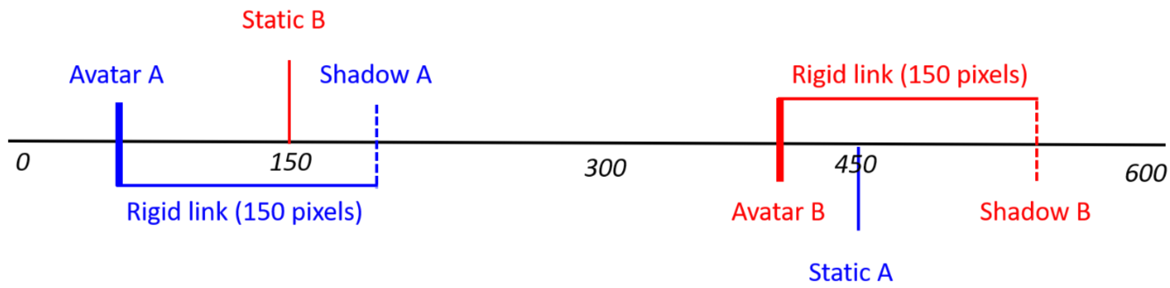

Figure 1. The virtual one-dimensional space with connected endpoints and the relation between entities; avatar A and B representing player A and B, respectively; shadow A and B representing the shadow of avatar A and B, respectively; and each player's static object.

2.3 Procedure

The total duration of the experiment was fifteen to twenty minutes. The first five minutes were used for instruction (see supplementary material B). The experiment consisted of six or ten one-minute rounds in which participants tried to complete the task of finding each other in the virtual space, using tactile feedback only. They were instructed to press a button (i.e. click) with their free hand at the moment that they were most confident of crossing the other avatar. Participants could use the entire minute to explore the virtual space and could also choose not to click in case they did not find the other avatar. They were instructed to stay in the other avatar's proximity after they clicked in order to help the other to complete the task in this cooperative game. Each new round started with random starting positions for both avatars. In order to find each other, participants were expected to distinguish between inanimate and animate entities (static object vs. shadow and other avatar), and between non-reactive and reactive entities (static object and shadow vs. other avatar). This was not specified in the instruction to the participants. Each round was followed by three self-report items on a tablet about participants' subjective experience of interaction during that previous round.

Participants did not receive any feedback on the behavior or clicking of their partner or on their own performance during the experiment. They were debriefed about the purpose of the experiment, and the strategy they could have used, after they finished the experiment and only if they were interested to know about it. Figure 2 shows an illustrative round.

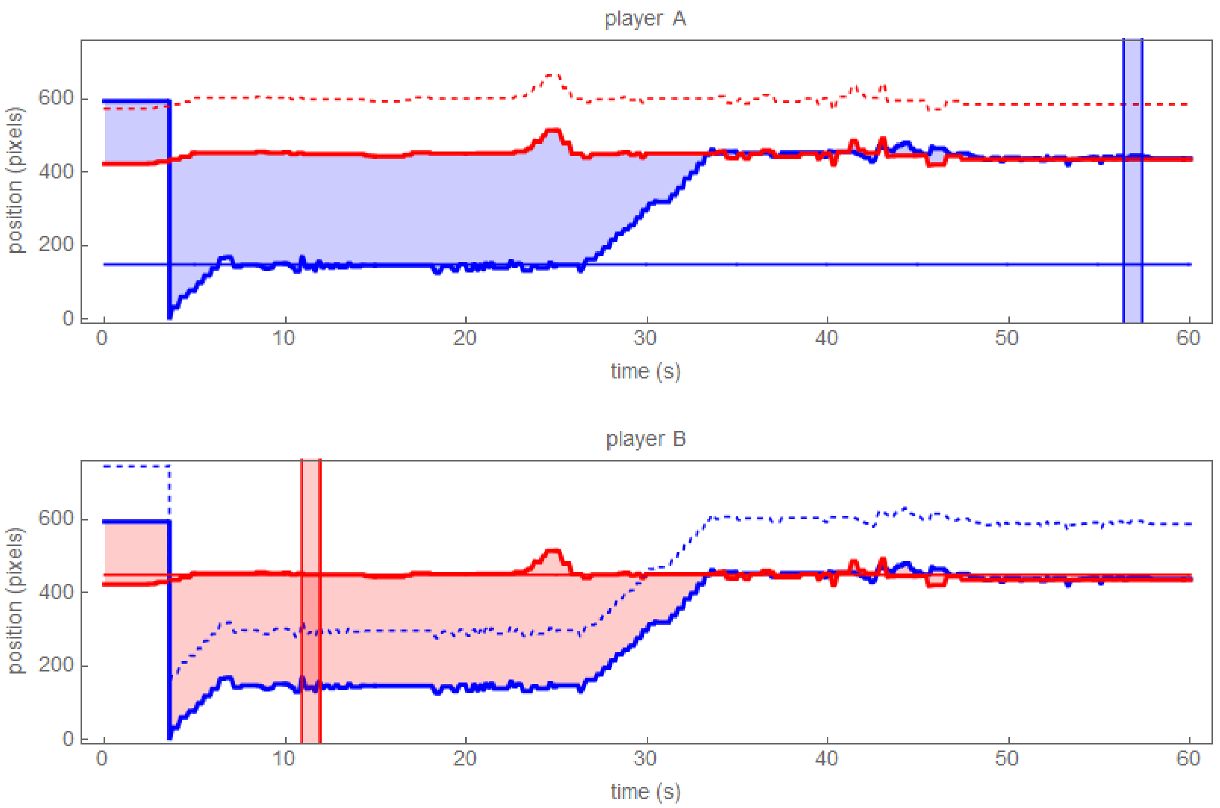

Figure 2. Recording of an illustrative round in which a pair (blue player A and red player B) interacted across the virtual one-dimensional space (y-axis) over the 60 seconds (x-axis), represented for each player separately. Solid bold blue and red lines represent the positions of the two players' avatars. Dotted blue and red lines represent the positions of their shadows, illustrated in the upper panel in red and in the lower panel in blue. Solid light blue and red lines represent the location of each player's static object. The vertical blue and red lines show the clicks. Blue player A clicked correctly (click assigned to the other avatar) as shown in the upper panel, while red player B clicked incorrectly (click assigned to the static object), as shown in the lower panel. Both players spent time exploring their respective static object. The red player remained with the static object, while the blue player started interacting with the red player after approximately 30 seconds.

2.4 Measures

2.4.1 Amount of time spent together

For each entity in the virtual space (i.e. other avatar, shadow, static object), the amount of time spent with this entity was computed as the total time (in steps of 10 milliseconds) during which the distance between the entity and the participant's avatar was below a given threshold. We set the threshold value at 70 pixels (see Froese et al., 2014). The amount of time spent together with the other avatar was defined as the time (in milliseconds) that the distance between the two participants' avatars was below 70 pixels (referred to as time spent together from here).

2.4.2 Correct detection

Both participants within a pair clicked independently and maximally once per round. The click was assigned to the entity closest to the participant's avatar within a distance of 70 pixels within one second before the click. This could be either the other avatar, the shadow, or the static object. If none of these entities were within the 70 pixels distance, the click was categorized as unclassified. Correct detection was defined as a click within 70 pixels distance from the other avatar. This is a binary variable that was calculated per round per individual. It was rated with '1' if a click was correct (i.e. assigned to the other avatar), and '0' if a click was incorrect (i.e. all the other instances where there was a click). Correct detection of the other is a variable that is independent of the other participant's click and will be referred to as correct detection from here.

2.4.3 Subjective experience of interaction

In order to measure participants' subjective experience of interaction, three items were used: '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?' Subjective experience of interaction was measured with a 7-point Likert scale ranging from '1' not at all to '7' very much. The items on the subjective experience of interaction were presented after each round to assess participants' experience during that entire previous round. The item 'How confident were you that you clicked correctly?' rated the confidence of clicks and included 'I haven't clicked' as an answer option, which was coded as a missing value. Subjective experience of interaction will be computed as the average of three items after rounds including a click, and as the average of two items after rounds without a click. This variable will be referred to as subjective experience from here.

2.4.4 Click success

We defined the variable Click success per round with four levels per pair. This variable was derived from correct detection, but is a paired variable where the value of the individual is dependent on the value of the other within the pair. It was coded as follows: '0' = no success (both players scored 0 on correct detection), '1' = single success (this player scored 1 on correct detection, the other scored 0), '2' = double success (both players scored 1 on correct detection within the same round, irrespective of the time interval within the round), and '3' = joint success (both players scored 1 on correct detection within a distance of 70 pixels within the same one second time interval). No click was coded as a missing value.

2.5 Analyses

For each analysis, age and gender were added as *a priori* covariates. Further, school, pair, and participant were added as levels in the multilevel analyses to account for the nesting of rounds within participants within pairs and within schools. If data were collapsed across rounds per individual, the participant level was left out. An exploratory factor analysis was conducted to statistically test whether the three items assessing subjective experience could be reduced to one or two variables.

In order to compare the mean amount of time spent together (i.e. with the other avatar) with the mean time spent with both the shadow and the static object, paired t-tests on collapsed data per individual across all rounds were used (hypothesis 1a). Logistic mixed-effect regression with only an intercept was conducted to test if the intercept was equal to zero, i.e. testing if the probability of correct detection was at chance level (0.5) (hypothesis 1b). To test the hypotheses that subjective experience (dependent variable) was related to time spent together (1c), proportion of correct detection (1d), or click success (1e), multilevel mixed-effect regression analyses were estimated in three separate models. For hypotheses 1c and 1d, data were collapsed per individual across all rounds. Proportion of correct detection was calculated per individual by dividing the total number of correct clicks by the number of total clicks.

Multilevel mixed-effect (logistic) regression analyses were fitted to examine whether the three main outcome variables time spent together, correct detection, and subjective experience were predicted by round as independent variable (hypotheses 2a, b and c). Random intercept and slope were allowed and only linear models were fitted.

Finally, to examine whether there was a difference in the average level of main outcome variables (variables time spent together, correct detection, and subjective experience) in round one to six versus round seven to ten, a dummy level (round one to six = 0; round 7 to 10=1) was used as predictor variable (hypotheses 3a, b and c). For these analyses, only the data from a subsample of 32 participants who completed ten consecutive rounds was used. All analyses were preregistered (confirmatory).

3. Results

3.1 Sample and data characteristics

3.1.1 Descriptives

The initial sample included 164 participants. Sixteen participants were excluded from analyses because of technical issues with the apparatus. The final sample included a total of 148 participants of whom 116 completed six rounds and 32 completed ten rounds of the PCE. Forty participants were attending 1st year, 32 3rd year, and 76 5th year in the secondary education system in Belgium. The age ranged from 12 to 19 years.

Visual inspection of the outcome variables and Shapiro-Wilk tests showed right-skewed distributions of the variable time spent together. In order to meet the requirement for normal distribution of outcome variables, the variables reflecting time spent together (as well as time spent with other shadow and static object) were square-root transformed. This transformation was selected because the data were right skewed and included zero values. In reporting the results, the mean values were back-transformed by squaring the values. To facilitate interpretation, we reported time spent together in seconds.

3.1.2 Exploratory factor analysis on subjective experience of interaction

Across rounds and per round, the three items used in the current study showed a significant Bartlett's test and a KMO above .5, fulfilling the requirements to conduct a factor analysis. The results indicated presence of one underlying factor. Our subjective experience score therefore reflected the explicit awareness of the other and the other's conscious awareness of the self, in combination with confidence of the presence of the other. Across rounds and per round, the mean score subjective experience indeed supported one underlying factor with an eigenvalue above 1. Mean scores of each item showed low uniqueness values (ranging from .10 to .39), indicating that their variance was well explained by the variable subjective experience. The inter-item reliability of the three items per round was high, with a Cronbach's alpha ranging from .85 to .92.

3.2 Social contingency detection across rounds

Time spent together - Across rounds, time spent together (mean=20s) was significantly higher than time spent with the shadow (mean=11s; $p<.001$) and the static object (mean=18s; $p=.002$). Time spent with the shadow was significantly lower than time spent with both the other avatar and the static object ($p<.001$).

Correct detection - Correct detection of the other was not at chance level ($p=.010$). Participants clicked during 79.2% of rounds (805 out of 1016 potential clicks). In 41.9% of these cases, the click was correct, i.e. detection of the other was successful. In 6.6% of cases, both the other avatar and either the other avatar's shadow or the static object were within 70 pixels of the avatar. In these cases, a click was assigned to the other avatar, i.e. defined as correct. Clicks were assigned to the static object in 33.8% of total clicks and to the shadow in 15.3% of total clicks. In 9.1% of cases, the clicks were categorized as unclassified because these did not occur within 70 pixels distance from any of the entities.

Subjective experience - Subjective experience increased when more time was spent together, although this did not reach statistical significance (B=.04 (.02), *95% CI*: -.00 to .08, p=.074) whereas subjective experience was significantly associated with click success (B=.18 (.04), *95% CI*: .10 to .27, p<.001). It was not associated with proportion of correct detection (B=-.01 (.41), *95% CI*: -.8 to .79, p=.973).

Click success (associated with subjective experience) - Across rounds, 58.1% of clicks were incorrect, 22.4% were single successes, 16.6% were double successes, and 2.9% were joint successes. Due to the low frequency of joint successes, these were counted as double successes, resulting in 19.5% of clicks falling into this category. Compared with an incorrect click, double success was associated with a significant increase of subjective experience (*B*=.23 (.07), *95% CI*: .10 to .36, *p*=.001). The difference between single success and double success was also associated with a significant increase of subjective experience (*B*=.22 (.08), *95% CI*: .07 to .38, *p*=.004). The difference between incorrect clicks and single success clicks was not associated with an increase of subjective experience (*B*=.01 (.06), *95% CI*: -.12 to .13, *p*=.929).

3.3 Learning of social contingency detection across rounds

For each outcome variable, we tested whether the average level changed throughout the experiment, from round one to six. First, round was not significantly associated with time spent together (*p*=.719), such that time spent together remained at a similar level during the experiment. This is illustrated in Figure 3, in addition to showing that the time spent with the shadow remained at a similar (lower) level. Moreover, the illustration shows a decreasing trend of time spent with the static object after the third round. Second, round was associated with a significant increase of probability of correct detection (*B*=.07 (.04), *95% CI*: 0.00 to .14, *p*=.05), such that there is some indication that the probability of correct detection increased per round. Click assignment started in the first round more or less at random with about the same number of clicks assigned to the other avatar, static object, or shadow. With successive rounds, there was a clear increasing trend, with half of the clicks assigned to the other avatar at the sixth round. The other half of click assignment was distributed over the other entities with decreasing numbers to the static object and the shadow. Correct detection was 25% after round one and 50% after round six. Lastly, round was associated with subjective experience (*B*=.07 (.01), *95% CI*: .04 to .10, $p<.001$), such that subjective experience increased across rounds.

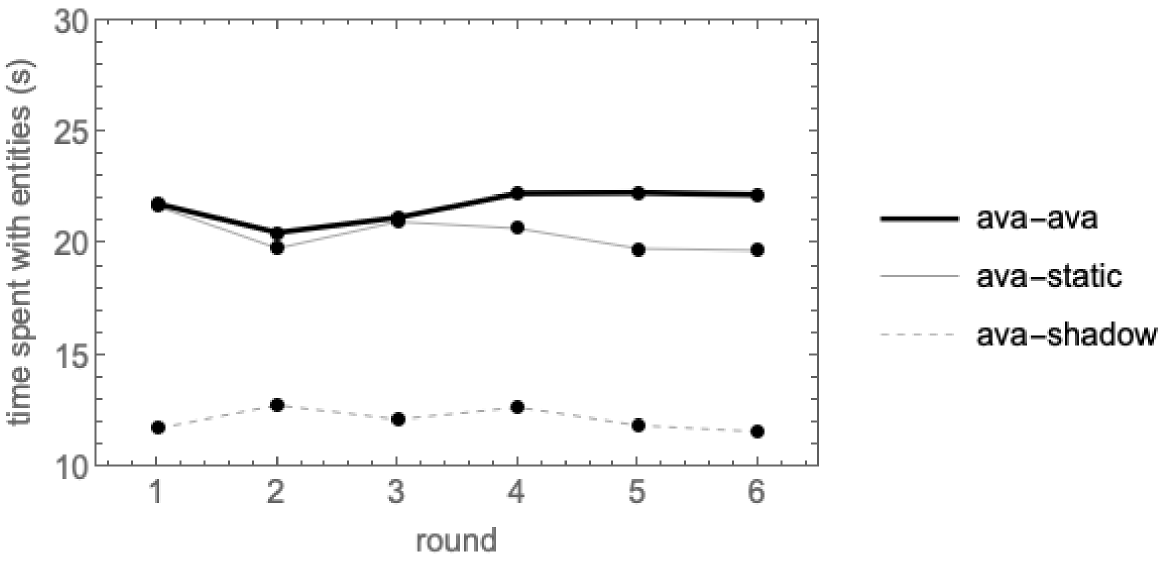

Figure 3. The amount of time spent with entities per round. The bold line represents the amount of time spent with the other avatar (ava), the light line represents the amount of time spent with the static object (static), and the dotted line represents the amount of time spent with the shadow (shadow).

3.4 Comparison of average social contingency detection levels between six-round version and extended ten-round version

In order to compare the six-round version with a ten-round version, a subsample of 32 participants who performed ten rounds was used for the analysis. The dummy variable reflecting either round one to six (0) or round seven to ten (1) was not significant in any of the associations tested for research question 2. This indicated that there was no evidence for a change in average level of time spent together, correct detection, and subjective experience in round one to six compared with the average level in round seven to ten. The results for time spent together and correct detection are illustrated in Figure 4, in which the average for these variables is shown per round.

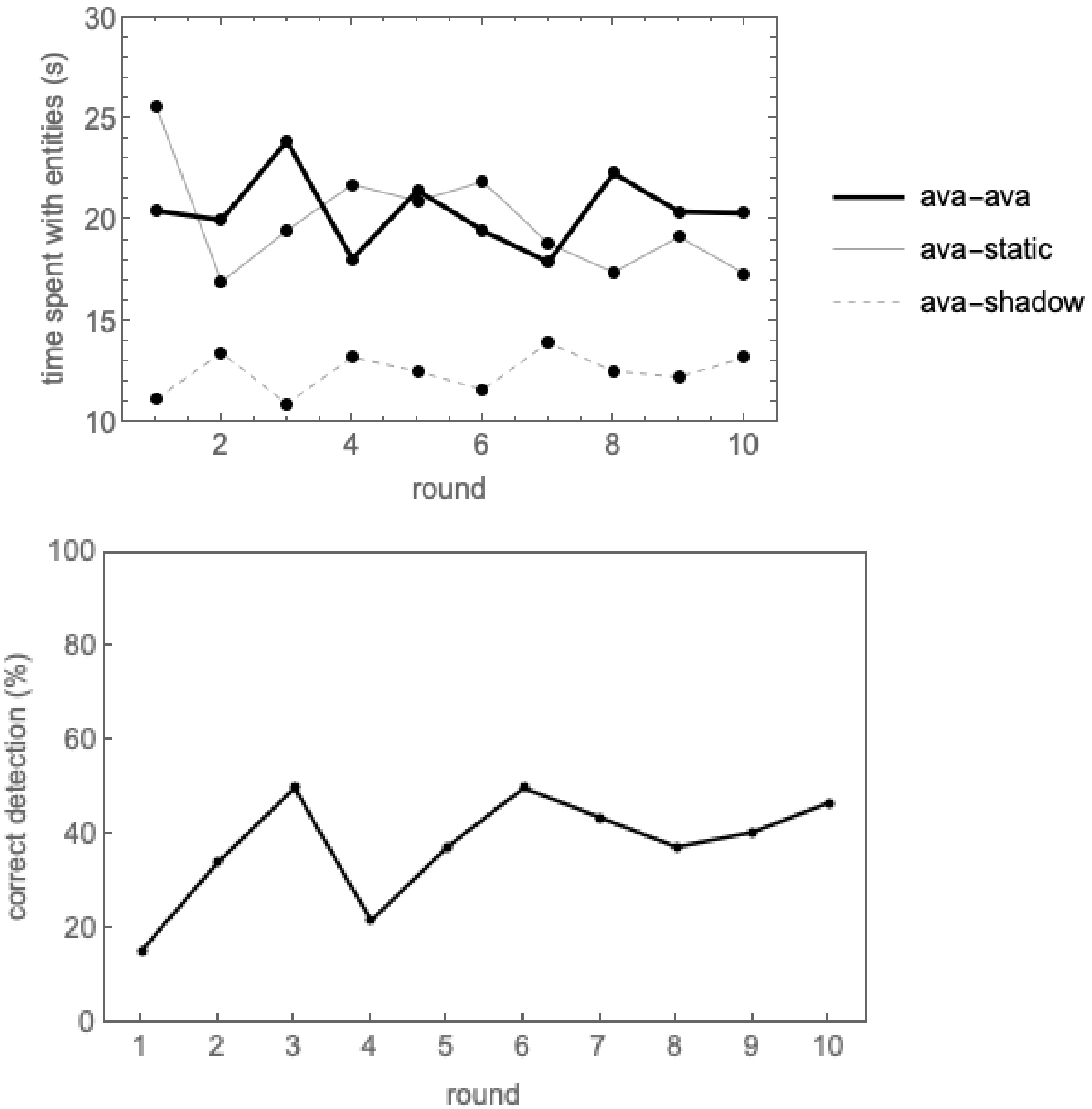

Figure 4. The average time spent with entities in seconds (upper panel) and the average correct detection in percentage of total clicks (lower panel) per round for the subsample of 32 participants who completed ten consecutive rounds. In the upper panel, the bold line represents the amount of time spent with the other avatar (ava), the light line represents the amount of time spent with the static object (static), and the dotted line represents the amount of time spent with the shadow (shadow).

3.5 Covariates

Age was a significant covariate in testing the association between round and correct detection, such that there is a significant positive effect on the average level of round and correct detection. Gender was not significant in any of the associations.

4. Discussion

4.1 Main findings

This is the first study that used the PCE in adolescents in order to assess real-time social contingency. Our results showed that the six-round version of the PCE had the capacity to assess social contingency detection in adolescents, across all rounds, in terms of amount of time spent together and correct detection of the other. Across rounds, correct detection of the other improved and the level of subjective experience of interaction increased. Importantly, we found subjective experience to be increased for rounds with double correct clicks compared to rounds with single and incorrect clicks. The average level of social contingency detection did not significantly change in round seven to ten compared with round one to six.

4.2 Comparison with previous findings

Overall, our six-round setup in adolescents has shown a similar capacity to assess social contingency detection as was shown in previous studies in adults that used a more extended setup (e.g. Auvray et al., 2009; Froese et al., 2014). This indicates that the setup used in this study is feasible in an adolescent population, and that the shortened version has a similar capacity to assess social contingency detection as did the longer version used previously. We have shown that correct detection of the other was, on average, not at chance level. Further, we reported a similar percentage of absent clicks compared with Froese et al. (2014). Contrastingly, differences between number of clicks assigned to the other avatar and the static object were less marked compared with this previous adult study. This may be explained by the current setup's absence of training rounds, which were included by Froese et al. (2014). In this training phase, participants became familiar with distinguishing the regular stimulation received while moving back and forth across a static object, and the comparatively regular stimulation received when two players engaged in a coordinated back-and-forth interaction (Di Paolo, Rohde, & Iizuka, 2008). Indeed, Figure 3 indicated that players started spending less time with the static object after three rounds, suggesting that they distinguished this entity from the other avatar after having experienced both stimulations in the first three rounds. Further, we reported a lower correct detection rate compared with Froese et al. (2014). This difference could be interpreted in light of a more advanced level of decision-making in adults, specifically affecting the decision to click. That is, although adolescents' number of explicit judgments about an interaction (i.e. clicking) was lower than what was found in adult studies, they did spent most time together and less time with other entities, which was also expected based on these previous adult studies (Auvray et al., 2009; Froese et al., 2014). They were also successful in ignoring the shadow, as evidenced by spending least time with this entity, most likely because of its unstable, non-responsive character, which did not need sustained attention to successfully reach the goal of the task. In other words, while participants were successful at engaging in interaction, they did not make this explicit as often as adults did. Although we cannot conclude from our findings whether this is due to being unaware of the other or lacking judgment while being aware of the other, we argue that this difference is likely to be explained by the age difference between the compared samples. Indeed, age had a significant positive effect on the average level of round and correct detection. This warrants future subgroup analyses of age, for instance to test the hypothesis that the capacity of making explicit judgments about social contingency continues to develop during adolescence.

This is the first study that replicated the increase of subjective experience in cases of mutual correct detection compared with single detection and incorrect detection, as was found by Froese et al. (2014). This serves as proof of principle that the subjective experience of social interaction is not something specific for one individual in the interaction or related to social cognitive capacities of one individual, but rather comes about as the result of a dynamic coupling of two individuals in the interaction. The partners in the dynamical system experience most interaction when both have detected the social contingency. This was further supported by the different associations between subjective experience and, on the one hand, proportion of correct detection (individual variable), and, on the other hand, click success (paired variable). As the three items assessing subjective experience formed one single factor, this indicated that participants were particularly aware of the other participant via the other's interactional directedness toward themselves. Taken together, these results demonstrate the importance of studying social interactive capacity for social contingency detection that is associated with the experience of interaction, rather than studying cognitive processes internal to the individual's brain (Buzan, Kupfer, Eastridge, & Lema-Hincapie, 2014).

Our findings in random pairs from the general population showed that time spent together did not change per round, indicating that participants kept exploring the space rather than increasingly staying with the other avatar. In contrast, Zapata-Fonseca et al. (2018) found controls to decrease their exploring behavior in interaction with individuals with autism spectrum disorder. This may also be explained by the difference in age, with an adult population more easily reaching a decision and sticking to what they think is the other person. Alternatively, it may be due to the characterization, with the healthy controls adapting their interaction strategy to their partner with autism spectrum disorder.

4.3 Are six rounds sufficient to capture social contingencies with the PCE in adolescents?

As illustrated in Figure 4, the level of time spent together, correct detection, and subjective experience did not further increase after six rounds when the experiment was extended with four additional rounds. After four rounds, there is a decrease in percentage of correct detection. This sudden drop to nearly the participants' average starting level suggests that something changed in explicitly judging about the interaction after a few rounds. This could be explained by the concept of (reinforcement) learning, including implicit and explicit learning (e.g. Barch et al., 2017; Berridge, 2004). The literature on sensorimotor learning in specific has suggested that this starts with implicit learning, followed by explicit learning (Taylor & Ivry, 2011; Taylor, Krakauer, & Ivry, 2014). These previous studies showed a decrease in performance when participants started to employ an explicit strategy to reach a goal, and it was suggested that this is due to a shift from action based on sensory-prediction error (i.e. difference between actual and predicted outcome) to action based on target error (i.e. difference between actual and targeted outcome). The latter can be interpreted as a shift to problem solving, in which participants attempt to use a cognitive strategy, which first leads to worse performance but is followed by a synergy of both ways of learning. This idea is in line with our findings in showing an increase of performance again after four rounds. It is also a hypothesis that requires further study, as the performance stabilized at a similar level as before, which might be lower than expected from a synergistic mode of sensitivity to social contingency detection. Nevertheless, this stabilization of performance did provide evidence that six rounds are sufficient to capture a stable level of social contingency detection and learning thereof. Our findings are also in concordance with subjective free-text reports obtained within the fifteen-round version by Froese et al. (2014), suggesting that players became aware of the other after only a few rounds already. More variation in time spent with entities in Figure 4 compared with Figure 3 is probably due to the smaller sample size used in the analysis of the ten-round version of the experiment. Indeed, the standard error for the mean values given in Figure 3 was lower than the standard error for the mean values given in Figure 4 (supplementary C). This was the case for each round, except for the fourth round, in which the standard deviation and error were higher in Figure 3 compared with Figure 4. The drop in time spent together during this round is in line with the sudden drop of proportion of correct detection during this round, potentially explained by the earlier mentioned concept of (reinforcement) learning. Overall, based on these results, a six-round version of the PCE seems reliable and valid in an adolescent sample. It would therefore be interesting to use this setup in order to further investigate the reason underlying the lower correct detection rate in adolescents compared with adults.

4.4 Future considerations regarding methodology

The 70-pixel interval used for click assignment could be tailored to the data to determine the specific sample’s optimal proximity range. This may be important in samples characterized by different styles of decision-making compared with healthy adults, such as in patients with psychosis (Garety et al., 2018), or in a younger population such as the sample used in the current study (Crone, 2013). Another consideration is to include measures of mutual coordination, for example by using complexity matching at the pair level (Kojima et al., 2017; Zapata-Fonseca et al., 2019), or time series analysis for turn-taking (Zapata-Fonseca et al., 2016). While a correct click was defined as the correct, but explicit, detection of the other from the experimenter’s point of view, the actual interaction or co-regulation might not always need to be made explicit in order to be successful from the participants’ subjective perspective. Indeed, our correct detection rate in adolescents was lower compared with adults. Further, Zapata-Fonseca et al. (2018) have shown that click correctness did not distinguish participants with high functioning autism from controls, while interaction patterns differed. Potential ambiguity in the interpretation of quantitative findings could be solved by including a qualitative aspect and compare this with the quantitative findings (Froese et al., 2014). Finally, we expect the PCE to explain more variance in social interaction compared with less ecologically valid experiments that focus on the individual. This hypothesis needs to be substantiated by first studying associations between our experimental findings and other ways of measuring social interaction, such as retrospective self-report questionnaires and momentary assessments in daily life. This would also provide studies investigating social interaction and social functioning with a paradigm to answer research questions about underlying mechanisms of social behavior, its development, and its potential variability within and between individuals.

5. Conclusion

The current findings indicate that the assessment and learning of social contingency detection can be achieved in an adolescent population by using a short and simple setup, without requiring training or complicated instructions. The potential role of age on social contingency detection warrants its inclusion in prospective studies that will aid in elucidating the complex nature of social interaction, even more if the link with social functioning can be established.

Open Practices Statements

None of the data or materials for the experiments reported here is available. The study was preregistered at the website of the Open Science Framework, available via https://osf.io/jmbdr/?view_only=9206a27ca3834a7da8da116b6154d1ad. Discrepancies between the preregistration and the final report are detailed in supplementary material D.

Funding

This work was supported by the FWO Odysseus grant to Inez Myin-Germeys (ZKD1475-03-W01). RF and LZ-F acknowledge financial support from the Universidad Nacional Autónoma de México (UNAM) through the DGAPA-PAPIIT project IA102619, and from the Consejo Nacional de Ciencia y Tecnología (CONACyT) through project Fronteras de la Ciencia 2016-01-2277. LZ-F is grateful for the scholarship 638215 granted by the CONACyT, and and a scholarship in the programme Research Grants-Bi-nationally Supervised Doctoral Degrees, 2019/20 granted by the German Academic Exchange Service (DAAD, Deutscher Akademischer Austauschdienst). RF is grateful for the Newton Advanced Fellowship awarded by the Academy of Medical Sciences through the UK Government.

References

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

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

Barch, D. M., Carter, C. S., Gold, J. M., Johnson, S. L., Kring, A. M., MacDonald, A. W., . . . Strauss, M. E. (2017). Explicit and implicit reinforcement learning across the psychosis spectrum. *Journal of Abnormal Psychology, 126*(5), 694-711. doi:10.1037/abn0000259

Barch, D. M., Pagliaccio, D., & Luking, K. (2016). Mechanisms Underlying Motivational Deficits in Psychopathology: Similarities and Differences in Depression and Schizophrenia. *Current Topics in Behavioral Neurosciences, 27*, 411-449. doi:10.1007/7854_2015_376

Berridge, K. C. (2004). Motivation concepts in behavioral neuroscience. *Physiology and Behavior, 81*(2), 179-209. doi:10.1016/j.physbeh.2004.02.004

Blakemore, S. J., & Choudhury, S. (2006). Development of the adolescent brain: implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry, 47*(3-4), 296-312. doi:DOI 10.1111/j.1469-7610.2006.01611.x

Bosco, F. M., Gabbatore, I., & Tirassa, M. (2014). A broad assessment of theory of mind in adolescence: The complexity of mindreading. *Consciousness and Cognition, 24*, 84-97. doi:10.1016/j.concog.2014.01.003

Brizio, A., Gabbatore, I., Tirassa, M., & Bosco, F. M. (2015). No more a child, not yet an adult": studying social cognition in adolescence. *Frontiers in Psychology, 6*. doi:ARTN 1011

10.3389/fpsyg.2015.01011

Buzan, R. D., Kupfer, J., Eastridge, D., & Lema-Hincapie, A. (2014). Philosophy of mind: Coming to terms with traumatic brain injury. *NeuroRehabilitation, 34*(4), 601-611. doi:10.3233/Nre-141071

Crone, E. A. (2013). Considerations of Fairness in the Adolescent Brain. *Child Development Perspectives, 7*(2), 97-103. doi:10.1111/cdep.12022

De Jaegher, H., & Di Paolo, E. (2008). Making Sense in Participation: An Enactive Approach to Social Cognition. In F. M. e. al. (Ed.), *Enacting Intersubjectivity*: IOS Press.

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

Deschamps, L., Lenay, C., Rovira, K., Le Bihan, G., & Aubert, D. (2016). Joint Perception of a Shared Object: A Minimalist Perceptual Crossing Experiment. *Frontiers in Psychology, 7,* 1059. doi:10.3389/fpsyg.2016.01059

Di Paolo, E. A., Rohde, M., & Iizuka, H. (2008). Sensitivity to social contingency or stability of interaction? Modelling the dynamics of perceptual crossing. *New Ideas in Psychology, 26*(2), 278-294. doi:10.1016/j.newideapsych.2007.07.006

Fett, A. K. J., Viechtbauer, W., Dominguez, M. d. G., Penn, D. L., van Os, J., & Krabbendam, L. (2011). The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: A meta-analysis. In (Vol. 35, pp. 573-588).

Froese, T. (2018). Searching for the conditions of genuine intersubjectivity: From agent-based models to perceptual crossing experiments. In N. A., L. de Bruin, & S. E. Gallagher (Eds.), *The Oxford Handbook of 4E Cognition*. Oxford: Oxford University Press.

Froese, T., & Di Paolo, E. A. (2011). The enactive approach Theoretical sketches from cell to society. *Pragmatics & Cognition, 19*(1), 1-36. doi:10.1075/pc.19.1.01fro

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

Garety, P., Freeman, D., Jolley, S., Ross, K., Waller, H., & Dunn, G. (2018). Jumping to conclusions: the psychology of delusional reasoning. *Advances in Psychiatric Treatment, 17*(05), 332-339. doi:10.1192/apt.bp.109.007104

Hermans, K. S. F. M., Achterhof, R., Myin-Germeys, I., Kasanova, Z., Kirtley, O., & Schneider, M. (2019). Improving ecological validity in research on social cognition. In E. Lewandowski & A. M. Moustafa (Eds.), *Social Cognition in Psychosis*: Elsevier Publishers.

Hutto, D. D., Herschbach, M., & Southgate, V. (2011). Editorial: Social Cognition: Mindreading and Alternatives. *Review of Philosophy and Psychology, 2*(3), 375-395. doi:10.1007/s13164-011-0073-0

Kelleher, I., Connor, D., Clarke, M. C., Devlin, N., Harley, M., & Cannon, M. (2012). Prevalence of psychotic symptoms in childhood and adolescence: a systematic review and meta-analysis of population-based studies. *Psychological Medicine, 42*(9), 1857-1863. doi:10.1017/S0033291711002960

Kirtley, O. J., Achterhof, R., Hagemann, N., Hermans, K. S. F. M., Hiekkaranta, A. P., Lecei, A., . . . Myin-Germeys, I. (in preparation). SIGMA, a longitudinal study of adolescent mental health and development using the experience sampling method: Cohort characteristics and protocol.

Kojima, H., Froese, T., Oka, M., Iizuka, H., & Ikegami, T. (2017). A Sensorimotor Signature of the Transition to Conscious Social Perception: Co-regulation of Active and Passive Touch. *Frontiers in Psychology, 8*, 1778. doi:10.3389/fpsyg.2017.01778

Murray, L., & Trevarthen, C. (1985). Emotional regulation of interactions between two-month-olds and their mothers. In T. M. Field & N. A. Fox (Eds.), *Social Perception in Infants* (pp. 177-197). Norwood, NJ: Ablex Publishers.

Nadel, J., Carchon, I., Kervella, C., Marcelli, D., & Reserbat-Plantey, D. (1999). Expectancies for social contingency in 2-month-olds. *Developmental science, 2*(2), 164-173. doi:10.1111/1467-7687.00065

Penn, D. L., Sanna, L. J., & Roberts, D. L. (2008). Social cognition in schizophrenia: an overview. *Schizophrenia Bulletin, 34*(3), 408-411. doi:10.1093/schbul/sbn014

Reddy, V. (2010). *How Infants Know Minds*: Harvard University Press.

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

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences, 36*(04), 393-414. doi:10.1017/S0140525X12000660

Schneider, M., Myin, E., & Myin-Germeys, I. (2019). Is theory of mind a prerequisite for social interactions? A study in psychotic disorder. *Psychological Medicine*, 1-7. doi:https://doi.org/10.1017/S0033291719000540

Simons, C. J., Bartels-Velthuis, A. A., & Pijnenborg, G. H. (2016). Cognitive Performance and Long-Term Social Functioning in Psychotic Disorder: A Three-Year Follow-Up Study. *PloS One, 11*(4), e0151299. doi:10.1371/journal.pone.0151299

Smetana, J. G. (2010). *Adolescents, Families, and Social Development. How Teens Construct Their Worlds.*: Wiley-Blackwell.

Smetana, J. G., Campione-Barr, N., & Metzger, A. (2006). Adolescent development in interpersonal and societal contexts. *Annual Review of Psychology, 57*(Petersen 1988), 255-284. doi:10.1146/annurev.psych.57.102904.190124

Taylor, J. A., & Ivry, R. B. (2011). Flexible cognitive strategies during motor learning. *PLoS Computational Biology, 7*(3), e1001096. doi:10.1371/journal.pcbi.1001096

Taylor, J. A., Krakauer, J. W., & Ivry, R. B. (2014). Explicit and implicit contributions to learning in a sensorimotor adaptation task. *Journal of Neuroscience, 34*(8), 3023-3032. doi:10.1523/JNEUROSCI.3619-13.2014

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. doi:10.3389/fpsyg.2016.01940

Zapata-Fonseca, L., Dotov, D., Fossion, R., Froese, T., Schilbach, L., Vogeley, K., & Timmermans, B. (2019). Multi-Scale Coordination of Distinctive Movement Patterns During Embodied Interaction Between Adults With High-Functioning Autism and Neurotypicals. *Frontiers in Psychology, 9*. doi:10.3389/fpsyg.2018.02760

Zapata-Fonseca, L., Froese, T., Schilbach, L., Vogeley, K., & Timmermans, B. (2018). Sensitivity to Social Contingency in Adults with High-Functioning Autism during Computer-Mediated Embodied Interaction. *Behav Sci (Basel), 8*(2). doi:10.3390/bs8020022

Appendices

Supplementary A: Hypotheses

Research question 1

Hypothesis 1a: Time spent together (milliseconds spent together within 70 pixels apart) will be higher than time spent with the other's shadow or the static object

Hypothesis 1b: Correct detection (binary variable with 0 if incorrect click and 1 if correct click) will not be at chance level

Hypothesis 1c: Subjective experience of interaction (mean score of two or three items) will increase as a function of time spent together

Hypothesis 1d: Subjective experience of interaction will increase as a function of proportion of correct detection

Hypothesis 1e: Subjective experience of interaction will increase as a function of click success (paired variable derived from correct detection)

Research question 2

Hypothesis 2a: Time spent together will increase as a function of round

Hypothesis 2b: Proportion of correct detection will increase as a function of round

Hypothesis 2c: Subjective experience of interaction will increase as a function of round

Research question 3

Hypothesis 3a: The intercept of time spent together in round seven to ten will not significantly differ from the intercept in round one to six

Hypothesis 3b: The intercept of proportion of correct detection in round seven to ten will not significantly differ from the intercept in round one to six

Hypothesis 3c: The intercept of subjective experience of interaction in round seven to ten will not significantly differ from the intercept in round one to six

Supplementary B: Instruction with metaphor

We are going to play a kind of game and you will be playing together with someone else. You will not be competing against each other, but you will collaborate. You will play in these two couples (*point to show participants who will play with who, back to back*). Now listen carefully: I would like you to imagine that you and your partner are walking together in a long, dark corridor. The corridor is round (*make gestures*) so you can walk round endlessly. This is a virtual task so we will not be really walking, but you can 'walk' by moving the trackball with your dominant hand (*show this*), back and forth. You can therefore either walk forwards or backwards. You cannot walk to the side, as the corridor is very narrow. You will be walking there with your partner only, so the other couple with be in another corridor. The corridor is dark, so you cannot see each other. You are indeed sitting back to back. You also will not be able to hear each other as I will ask you in a minute to put on your headphones. Now, it is your task to find each other in the dark corridor. How can you achieve this, without seeing or hearing each other? (*Ask participants to answer*) Yes, by feeling each other. You will not do this in real life again, but you will 'feel' the other person's passing by receiving a vibration on your hand. You will just be walking through the corridor and each time you pass the other, you will feel this vibration. You have one minute to find each other, and you can see this minute counting back at your screen. At the moment you think you are crossing the other and you are most certain about, you can push the blue button with your other hand. You can only push this button once per minute, so you can use the entire minute to do so. If, at the end of the minute, you feel you did not find the other, you do not need to push the button. You you either do it once, or not at all. In order to make this task a bit harder, you can also encounter a chair and another moving object in the corridor, next to your partner. These two other objects will also give you a vibration on your hand if you cross them, similar to the vibration you receive when you cross your partner. Now it is up to you to find a way to find your partner, and you should collaborate in doing so. If you pushed the button because you feel you found your partner, you should continue playing until the minute ends in order to still help your partner to find you as well. We will play six rounds of one minute. Is your task clear? After each round, I will ask you to complete three questions on your tablet.

Supplementary C: Standard deviation and standard error per round as illustrated in Figure 3 and 4

Table 1. Standard deviation (std) and standard error (se) for time spent with the other avatar (ava-ava), time spent with the static object (ava-static), and time spent with the shadow (ava-shadow) per round. The mean values of these variables represent the lines in Figure 3 (N=116) and Figure 4 (N=32).

			Round									
			1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
N=116 (Figure 3)	*ava-ava*	*std*	9.94	7.91	8.71	28.19	9.39	10.09				
		se	0.92	0.73	0.81	2.62	0.87	0.94				
	ava-static	*std*	13.44	11.96	11.36	59.26	12.06	11.47				
		se	1.25	1.11	1.05	5.50	1.12	1.07				
	ava-shadow	*std*	6.47	7.09	5.63	19.68	6.09	5.88				
		se	0.60	0.66	0.52	1.83	0.57	0.55				
N=32 (Figure 4)	*ava-ava*	*std*	10.47	6.35	8.46	4.07	8.46	6.34	8.87	8.39	9.03	10.56
		se	1.85	1.12	1.50	0.72	1.50	1.12	1.57	1.48	1.60	1.87
	ava-static	*std*	13.75	8.79	12.48	9.65	13.45	12.69	11.03	9.83	12.31	11.95
		se	2.43	1.55	2.21	1.71	2.38	2.24	1.95	1.74	2.18	2.11
	ava-shadow	*std*	5.43	4.73	5.61	6.67	6.97	5.89	6.39	6.31	4.53	7.72
		se	0.96	0.84	0.99	1.18	1.23	1.04	1.13	1.12	0.80	1.37

Supplementary D: Transparent changes document

Study Information

Discrepancies between the preregistration of this study (published online 5-9-18: https://osf.io/jmbdr/?view_only=9206a27ca3834a7da8da116b6154d1ad) and the final report before submission were added in another font and color in every section. Overall, it became clear over the course of conducting the analyses and writing up the report that the final manuscript should be simplified and include a comprehensible overview of the main goals for a methodological paper. Based on this consideration, we decided to change variable names and leave out some more advanced or content analyses.

1. The Perceptual Crossing Experiment in adolescents

1.1. Running title: PCE in adolescents

2. Authorship

K.S.F.M. Hermans, L. Zapata-Fonseca, Z. Kasanova, R. Fossion, T. Froese, I. Myin-Germeys *(order to be established)*

Order changed to Karlijn S.F.M. Hermans, Zuzana Kasanova, Leonardo Zapata-Fonseca, Ginette Lafit, Ruben Fossion, Tom Froese, Inez Myin-Germeys

Ginette Lafit was included for statistical support.

3. Research Questions

1) Does the modified 6-round version of the Perceptual Crossing Experiment (PCE) assess social contingency detection (measured as time spent with other avatar, correct click proportion, and perceptual awareness) in adolescents across all rounds?
2) Does the modified 6-round version of the PCE detect gradual acquisition of social contingency detection per round?
3) In a modified 10-round version of the PCE, does the level of acquisition of social contingency detection change in round 7 to 10 compared to round 1 to 6?

1) Does the modified six-round version of the PCE assess overall social contingency detection measured as amount of time spent together, correct detection of the other, and subjective experience

2) Does the modified six-round version of the PCE assess learning of social contingency detection across the six rounds?

3) In a modified ten-round version of the PCE, does the average level of social contingency detection change in round seven to ten compared to round one to six?

The variable names to assess social contingency have changed for the sake of simplicity. This is further detailed later in the method section. The same applied to 'gradual acquisition' which was translated to 'learning'. Research question 3 has changed to only comparing average levels of outcome variables instead of learning of these variables. This was due to power issues with comparability analysis (N=32). The goal of the task was still reached, and this change of research question/hypothesis and reason for it was addressed in the limitation section of the manuscript.

4. Hypotheses

1) Does the modified 6-round version of the Perceptual Crossing Experiment (PCE) assess social contingency detection (measured as time spent with other avatar, correct click proportion, and perceptual awareness) in adolescents across all rounds?

- Time spent with the other avatar (measured in probability density function; PDF and cumulative density function; CDF) will be higher than time spent with other entities (shadow of the other avatar or static object)

Time spent together was translated to milliseconds as a temporal variable was easier to interpret than a spatiotemporal variable.

- The probability of a correct click (defined as a click assigned to the other avatar based on proximity) will be above chance level

We changed 'will be above chance level' to 'not at chance level' because this was more accurate in terms of what we could conclude from this test.

- Perceptual awareness (measured with 3 adapted items of the Perceptual Awareness Scale; PAS) will increase as a function of time spent with the other avatar
- Perceptual awareness will increase as a function of correct click proportion (defined as a click assigned to the other avatar based on proximity)
- Perceptual awareness will increase as a function of click success (dyadic score defined based on correct clicks within dyad)
- Perceptual awareness will correlate with strategy (measured with an open question and categorized into different levels)

We have changed the naming of ‘perceptual awareness’ to ‘subjective experience’. The last hypothesis was left out because it would go beyond the scope of the current manuscript goals to assess capacity of the PCE.

2) Does the modified 6-round version of the PCE detect gradual acquisition of social contingency detection per round?

- Time spent with the other avatar will increase as a function of round
- Proportion of correct click will increase as a function of round
- Perceptual awareness will increase as a function of round

3) In a modified 10-round version of the PCE, does the level of acquisition of social contingency detection change in round 7 to 10 compared to round 1 to 6?

- The intercept and slope of time spent with the other avatar in round 7 to 10 will not significantly differ from the intercept and slope in round 1 to 6
- The intercept and slope of correct click proportion in round 7 to 10 will not significantly differ from the intercept and slope in round 1 to 6
- The intercept and slope of perceptual awareness in round 7 to 10 will not significantly differ from the intercept and slope in round 1 to 6

As previously described for research question 3, we have left out testing for slope as the goal of this research question was also achieved by looking at average levels of outcome variables. Comparative analysis of slopes was not feasible with the sample size of the subsample.

Sampling Plan

5. Existing data

For this pre-registration, registration prior to analysis of the data is applicable. As of the date of submission, the data exist and the first two authors (K.H. and L.Z.) have accessed it, though no analysis has been conducted by K.H. related to the research plan (including calculation of summary statistics). She submitted the pre-registration and worked strictly separate from L.Z. and R.F., who already started pre-processing of the data.

As an addition, this amendment that describes discrepancies between the original preregistration and the final manuscript has been published before submission of the manuscript. Major changes have also been addressed in the manuscript itself.

6. Explanation of existing data

Our data have already been collected by K.H. who, therefore, accessed the data. However, the raw data cannot be interpreted without pre-processing steps, such as visualization of the data and computing variables to be used as outcome variables in the planned analysis described in this pre-registration. These pre-processing steps were solely performed by L.Z. and R.F. who have worked on this in a different location. K.H. has only seen a visualization produced from a random sample of the data (drawn by L.Z) which was used for an internal lab presentation as a visual example of how a social interaction would look in the data. L.Z. and R.F. will continue working on the pre-processing steps independently of K.H. until the study has been pre-registered on the OSF site. Only after pre-registration, will the actual analyses be performed to test the hypotheses. These will be confirmatory tests.

At this moment in time, we have no planned exploratory analyses. Anything that would warrant further exploration will be registered separately, before conducting analyses.
No exploratory analyses were conducted.

7. Data collection procedures

The data have already been collected between February 2018 and May 2018. The sample comprised adolescents, aged 12 to 16, from the general population in Dutch-speaking Belgium.
Correction: The age ranged from 12 to 19.
Participation in the PCE was part of participation in a longitudinal cohort study, for which each participant received a 10 euro voucher after completion. The exclusion criterion for the PCE was an inadequate level of Dutch or English and therefore failure to understand the instruction.

The total duration of the experiment was 20 minutes. The first five minutes were used for instruction. This was followed by completing two baseline questions on familiarity with the partner in the experiment. The experiment itself consisted of 6 or 10 rounds of one minute, each, followed by 3 self-report adapted PAS items (Ramsøy & Overgaard, 2004; Froese et al., 2014) on a tablet, using the data management software REDCap (http://project-redcap.org/). The experimental set-up was based on previous studies using the PCE (e.g. Froese et al., 2014) in which participants had to collaboratively interact with a partner within a minimalistic set-up, only relying on haptic feedback. Participants moved through a virtual one-dimensional space with connected endpoints and aimed to encounter the other avatar (partner), establishing an interaction. They could additionally encounter a static object and the shadow of the other avatar, which moved at a fixed distance of this avatar. Both avatars moved through space with a trackball controlled by their dominant hand and each encounter with any of the three entities (static, shadow avatar B, other avatar B) elicited haptic feedback by a vibration on the same hand. The movement and vibration on the same hand warrants the sensorimotor loop to be experienced as integrated and aimed to provide an embodied account of the interaction. The virtual loop consisted of 600 pixels and the entities (static, shadow and avatar) are 4 pixels wide. On the virtual line, the two static objects (one for each avatar) were fixed at 150 and 450 pixels. The distance between the avatar and the shadow was 150 pixels.

8. Sample size

The sample included 160 participants, of which 48 were in 1st grade of secondary school, 35 in 3rd grade, and 77 in 5th grade. Demographical data was missing for two participants. The data will be analyzed in dyads who performed the task together, so the sample included 80 dyads.

After a closer look at the data we had to correct this to the following: "The initial sample included 164 participants. Sixteen (four experiments) participants were excluded from analysis because of technical issues. The final sample included 148 participants of whom 32 completed ten rounds. The total sample included 40 participants in 1st grade, 32 in 3rd grade, and 76 in 5th grade." This correction was anticipated in this preregistration (see 22. Missing data).

9. Sample size rationale

For the sample size, we maximized the number of participants that could be tested within the larger project as part of which the data were collected. In each group of 20 to 24 participants in a test session, eight participants were randomly selected to perform the PCE. This sample size exceeded the sample sizes of previous studies using the PCE, which included 20 (Auvray et al,. 2009, Zapata-Fonseca et al., 2018) and 34 (Froese et al., 2014) participants.

Variables

10. Manipulated variables

Not applicable.

11. Measured variables

1. Spatiotemporal position of click

For each round, the position of one click was registered at a spatiotemporal point in time when the LED light of the switch (button) was turned on. Only the first click was registered.

This variable has been simplified for the report. It was translated to a temporal variable computed as the total time (in milliseconds) during which the distance between the entity and the participant's avatar was below a given threshold. This given threshold was dx=70 (pixels).

2. Spatiotemporal position avatar

For each avatar, the spatiotemporal position on the one-dimensional virtual line was registered with a sampling frequency of 10 Hz.

For the same reason as mentioned above, we left these technical details out of the final report.

3. Perceptual awareness: 3 adapted PAS items (Ramsøy & Overgaard, 2004; Froese et al., 2014)

The rationale for slight changes from Froese et al. (2014) will be discussed in the method section of the paper. After each round, participants answered three questions about their experience in previous round.

- To what extent did you feel that the other could sense your presence? *Not at all (1) – very much (7)*
- To what extent did you feel you were doing something together? *Not at all (1) – very much (7)*
- How confident were you that you clicked correctly? *Not at all (1) – very much (7) – I haven't clicked*

We emphasized that these items were adapted from the items used in the previous adult study conducted by Froese et al. (2014) instead of the paper that originally described the Perceptual Awareness Scale (Ramsøy & Overgaard, 2004).

4. Qualitative question about strategy

After completion of the experiment (6 or 10 rounds), participants answered two questions about the strategy they had used throughout the experiment to successfully find each other.

- Did you use a strategy? *Yes – No*
- If yes, describe the strategy you used

The open question has been coded by K.H. and will be coded by Z.K., as specified in section 18 “Transformations”.

This section was left out because it was considered beyond the scope of a purely methodological report. It will be described later in a short report. The rationale for investigating this association has been addressed in the discussion.

12. Indices

In order to compute the outcome variables, the optimal range of proximity (between the avatar and another entity) needs to be established. Due to a lack of comparable previous research in an adolescent population with the modified, shorter set-up, we will define the optimal range based on data-driven analysis as described in this section.

It was considered beyond the scope of this methodological paper to conduct the more advanced step of data-driven analysis before testing whether this would be worthwhile. The current analyses were therefore based on previous adult studies. The findings resulted in limitations and future suggestions that were added to the discussion section.

1. Assignment to entities based on proximity

For all time steps (100ms) of a round, the distance y=abs(x1-x2) is calculated for the position x1 of a player with the position x2 of each other entity the player can interact with (i.e., the other avatar, the shadow of the other avatar, or the static object), resulting in 5 different distance time series y(t). For each distance time series, a frequency function and both probability density function PDF(y) and cumulative density function CDF(y) can be calculated. The frequency function ranges from y=-300 (300 pixels to the left of the other entity) to y=300 (300 pixels to the right of the other entity. One can define a proximity parameter, yMax, the maximal distance that the player can be from another entity and still be considered “proximate”. As an initial hypothesis, we consider as “proximate” a distance of yMax=75, because the distance from a player to his shadow is 150 pixels, and for distances larger than yMax=75 pixels it would not be possible to decide whether a player is close to the other avatar or to the other’s shadow. The total time T spent “proximate” to an entity is the cumulative sum of the area (CDF) of the corresponding distance frequency function histogram from -yMax to +yMax, i.e. the total number of time steps that the player was within a distance of 75 pixels from the other entity. yMax=75 is an initial hypothesis. We can study the function T(yMax), i.e. how the time T spent "proximate" to another entity varies as a function of the maximal distance yMax. The optimal proximity parameter should be the maximal distance yMax that allows for optimal distinction between closeness of a player with the other avatar and closeness with shadows or static objects. Based on this data-driven analysis we will determine whether the initial hypothesis of yMax=75 is the optimal range of proximity, or whether another value should be chosen. If the latter is true, we will report this and compute the outcome variables based on this optimal value.

The PDF and CDF will be calculated for time spent with each entity, resulting in 2x4 variables: time spent with other avatar, time spent with other shadow, time spent with static object, time spent unrelated to any entity.

Before we started analyses, we decided to use dt=100 (units of 10ms, thus dt=1s) and dx=70 (pixels) based on the previous adult study by Froese et al. (2014) on which our setup was mainly based. We also translated the PDF and CDF to reflect milliseconds in which the avatar was within 70 pixels of a given entity.

2. Click assignment

Clicks given within a distance yMax from another entity will be assigned to that entity. If at the time of the click 2 or more entities are within the "proximity" range, we will determine which of the entities was in this "proximity" range for the longest time interval immediately before the click (up to 1 second before). The "proximity" range will be referred to as the optimal range.

For the sake of simplicity, we used a temporal variable and left out this 'proximity range'. We did not determine which of the entities was in the 'proximity range' longest but instead assigned the click to the entity closest to the avatar in case two entities were within range.

A correct click is defined as a click assigned to the other avatar within the optimal range. All other clicks will be defined as incorrect. No click will be defined as a missing value (9). The proportion of correct clicks will be defined as number of correct clicks within 6 rounds divided by the number of total clicks.

Click success will be coded as a dyadic score, based on correctness of clicks for each round:

0 – Click unrelated to any entity or click assigned to static object or shadow → incorrect click

1 – Individual click assigned to other avatar (correct click) → single success

+1 – Both individuals clicked within same optimal range (assigned to other avatar) → joint success

This can result in the following dyadic scores combinations: 0 (0-0), 1 (0-1 or 1-0), 1 (1-1 in separate optimal range), and 2 (1-1 in same optimal range).

We specified this original distinction more based on previous literature: "The four levels include '0' No success (both players scored 0 on 'Correct detection'), '1' Single success (this player scored 1 on Correct detection, the other scored 0), '2' Double success (both players scored 1 on Correct detection within the same round), and '3' Joint success (both players scored 1 on Correct detection within a distance of 70 pixels). No click was coded as a missing value."

After the analyses were conducted, the number of occurrences of Joint success were too low to compare with other levels. Therefore, we decided to combine Double and Joint success into the category Double success. We have reported this in the Result section.

3. Perceptual awareness

In order to measure perceptual awareness of the social interaction (after each round), three items adapted from the PAS (Ramsøy & Overgaard, 2004; Froese et al., 2014) will be used. We will conduct an exploratory factor analysis (EFA) to statistically test whether the three adapted PAS variables can be reduced to 1 or 2 variables. We will use a cut-off factor loading of 0.3 to decide if we combine items into one variable "Perceptual awareness". If the factor loading does not exceed the cut-off value, the adapted PAS items will be added as separate predictors and correction for multiple comparisons will be applied.

4. Strategy

An open question on strategy used in each round will be coded based on a coding scheme outlined in section 12 "Transformations". This will result in a variable "Strategy" with four levels ranging from no strategy to advanced strategy.

This section was left out for reasons mentioned earlier.

Design Plan

13. Study type

The study type is an experiment as we randomly assigned participants from classes to participate in the experiment. However, our experiment should not be viewed as an intervention, but rather as observation in a random sample, closely resembling a lab experiment.

14. Blinding

No blinding is involved in this study.

15. Study design

Our study design is a repeated measures multi-level design with subjects nested within dyads nested within schools.

16. Randomization

We have randomly assigned participants to perform the experiment by using a website (www.random.org) that generates random numbers from a range of numbers. We assigned participants to the experiment prior to the day of testing. If any assigned participants were absent on the day of testing, we selected the next person on the class list.

Analysis Plan

17. Statistical models

Research question 1

Within our dataset, missing data are potentially not missing at random (MNAR) due to a higher likelihood of no click in first rounds compared to later rounds. We will therefore perform a sensitivity analysis with plausible MNAR models and see how consistent the results are for the different models (Allison, 2014).

We have not conducted this sensitivity analysis with plausible MNAR models because no click was reported as such and provided valuable information in itself. For the sake of comparability, we have followed previous studies in this regard.

Data from 6 rounds (n=180) will be used for these analyses.

N=148. We had to leave out a number of participants because they did not complete six consecutive rounds. This was due to a change of design during the data collection phase, not due to drop-out. We did not anticipate this for the preregistration.

- *Time spent with the other avatar (measured in PDF and CDF) will be higher than time spent with other entities (shadow of the other avatar or static object)* – Perform repeated measures one-way ANOVA with "time spent with other avatar", "time spent with other shadow”, “time spent with static object”, and “time spent unrelated to any entity” (measured in PDF and CDF) as variables. Test if the means of time spent with each entity differ from each other. If applicable, follow up with a post-hoc test to determine which means differ from each other.

- Hypothesis 1a was tested by performing paired t-tests on collapsed data per individual across all rounds, comparing the mean amount of time spent together (i.e. with the other avatar) with the mean time spent with both the shadow and the static object.

We have collapsed across rounds here to retrieve a value for Time spent with other avatar (together), with other shadow, and with the static object. As the Time spent together was primary to the hypothesis, we performed several t-tests in order to compare these variables against the Time spent together. We could not run a repeated measures ANOVA here as we could not add levels (School, Pair) as random effects.

- *The probability of a correct click (defined as a click assigned to the other avatar based on proximity) will be above chance level* – Logistic regression for repeated measures with only an intercept and a random fixed effect. Test if the intercept is equal to zero. This is equivalent to test that the proportion of clicking correct or incorrect are equal (chance performance).

- For hypothesis 1b, logistic mixed-effect regression with only an intercept was conducted to test if the intercept was equal to zero, i.e. testing if the probability of Correct detection was at chance level (0.5).

We could only test the hypothesis if the probability was at chance level or not.

- *Perceptual awareness (measured with 3 adapted items of the PAS) will increase as a function of time spent with the other avatar (in terms of PDF and CDF)* – Perform multilevel mixed-effect regression analysis with “perceptual awareness” as dependent variable and “time spent with other avatar” as independent variable. Add “round” as a covariate to control for differences between rounds. Add random effects for “subject”, “dyad” and “school”.

- *Perceptual awareness (measured with 3 adapted items of the PAS) will increase as a function of correct click proportion (defined as a click assigned to the other avatar based on proximity)* – Perform multilevel mixed-effect regression analysis with “perceptual awareness” as dependent variable and “correct click proportion” (binary) as independent variable. Add “round” as a covariate to control for differences between rounds. Add random effects for “subject”, “dyad” and “school”.

Proportion of Correct detection was calculated per individual by dividing the total number of correct clicks by the number of total clicks (as described previously in the preregistration as well).

- *Perceptual awareness (measured with 3 adapted items of the PAS) will increase as a function of click success (dyadic score defined based on correct clicks within dyad)* – Perform multilevel mixed-effect regression analysis with "perceptual awareness" as dependent variable and "click success" (with 3 levels) as independent variable. Add "round" as a covariate to control for differences between rounds. Add random effects for "subject", "dyad" and "school".

- The hypotheses with Subjective experience as outcome variable (1c, 1d, 1e) were tested by performing multilevel mixed-effect regression analyses, for hypotheses 1c and 1d on collapsed data per individual across all rounds. Separate analyses were conducted with Subjective experience as dependent variable and Time spent together, proportion of Correct detection, and Click success as independent variable.

Instead of adding Round as a covariate, we collapsed across rounds for the hypotheses testing with Perceptual awareness (or in the final report 'Subjective experience') as outcome variable.

- *Perceptual awareness (measured with 3 adapted items of the PAS) will correlate with strategy (measured with an open question and categorized into different levels)* – Perform repeated measures correlation to determine the common within-individual association between "perceptual awareness" and "strategy".

This was taken out for the current manuscript as inclusion was beyond its scope.

<u>Research question 2</u>

For model selection, we will compare different model fit statistics such as the Aikaike Information Criterion (AIC) to select the best fitting model.

For sake of simplicity, for each model, random intercept and slope were allowed and only linear models were fitted.

Data from 6 rounds will be used (n=160).

N=148.

- *Time spent with the other avatar (measured in PDF and CDF) will increase as a function of round* – Perform multilevel mixed-effect regression analysis with "time spent with other avatar" as dependent variable and "round" (with 6 levels) as independent variable. Add random effects for "subject", "dyad" and "school". Compare model fit statistics for linear, quadratic, and cubic models.

- *Proportion of correct click (assigned to other avatar) will increase as a function of round* – Perform multilevel mixed-effect logistic regression analysis with "proportion of correct click" as dependent variable and "round" (with 6 levels) as independent variable. Add random effects for "subject", "dyad" and "school". Compare model fit statistics for linear, quadratic, and cubic models.

- *Perceptual awareness will increase as a function of round* – Perform multilevel mixed-effect regression analysis with "perceptual awareness" as dependent variable and "round" (with 6 levels) as independent variable. Add random effects for "subject", "dyad" and "school". Compare model fit statistics for linear, quadratic, and cubic models.

- Hypotheses 2a, 2b, and 2c were tested by performing multilevel mixed-effect (logistic) regression analyses with the three main outcome variables (i.e. Time spent together, Correct detection, and Subjective experience) as dependent variables and Round as independent variable.

We did not compare model fit statistics.

Research question 3

For model selection, we will compare different model fit statistics such as the Aikaike Information Criterion (AIC) to select the best fitting model.

We did not compare model fit statistics.

Create a dummy variable to compare round 1 to 6 (level 0) with round 7 to 10 (level 1). Data from 10 rounds will be used (n=32).

- *The intercept and slope of time spent with the other avatar (measured in PDF and CDF) in round 7 to 10 will not significantly differ from the intercept and slope in round 1 to 6* – Perform multi-level linear regression analysis with "time spent with other avatar" as dependent variable and "round" (with 10 levels) as independent variable. Add a dummy variable to compare intercept and slope of round 1 to 6 (dummy level 0) to round 7 to 10 (dummy level 1).

- *The intercept and slope of correct click proportion in round 7 to 10 will not significantly differ from the intercept and slope in round 1 to 6* – Perform multi-level linear regression analysis with "correct click proportion" as dependent variable and "round" (with 10 levels) as independent variable. Add a dummy variable to compare intercept and slope of round 1 to 6 (dummy level 0) to round 7 to 10 (dummy level 1).

- *The intercept and slope of perceptual awareness in round 7 to 10 will not significantly differ from the intercept and slope in round 1 to 6* – Perform multi-level linear regression analysis with "perceptual awareness" as dependent variable and "round" (with 10 levels) as independent variable. Add a dummy variable to compare intercept and slope of round 1 to 6 (dummy level 0) to round 7 to 10 (dummy level 1).

A dummy was indeed added but this only tested the average level of outcome variables in round 1-6 compared with round 7-10.

18. Transformations

Strategy – The following coding scheme was used to quantify the qualitative question on the strategy that was used.

0: No strategy (indicated that no strategy was used)

1: Unspecific and vague (e.g. just feeling)

2: Distinguished between static object and moving entity (e.g. giving meaning to perceived "longer" and "shorter" vibrations)

3: Distinguished between responsive and non-responsive moving entity (e.g. following other avatar, going back and forth, mentioning the other avatar)

The first coding has been performed by K.H., who also collected the data. Z.K. will apply the same coding system and interrater agreement will be calculated.

This section was left out because the related hypothesis was left out.

19. Follow-up analyses

Not applicable.

20. Inference criteria

We will be using one-tailed tests (for directional hypotheses) and two-tailed tests (for non-directional hypotheses) with a significance level of p=0.05 (specified in section 17 "Statistical models"). Corrections for multiple comparison will be conducted for analyses in which this is applicable.

21. Data exclusion

We excluded 15 participants for technical reasons and another 68 because they did not complete 6 rounds.

22. Missing data

We excluded participants who provided incomplete data from the analyses. The exact number with rationale will be specified in the method section of the paper.

Other