# MASTERARBEIT | MASTER'S THESIS

Titel | Title

Affiliative relationships in pre-school children: Using an automated tracking approach to test the explanatory power of simple spatial measures


verfasst von | submitted by

Růžena Ceral B.Sc. RWTH


angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien | Vienna, 2025

| | |
|---|---|
| Studienkennzahl lt. Studienblatt \| Degree programme code as it appears on the student record sheet: | UA 066 878 |
| Studienrichtung lt. Studienblatt \| Degree programme as it appears on the student record sheet: | Masterstudium Verhaltens-, Neuro- und Kognitionsbiologie |
| Betreut von \| Supervisor: | Univ.-Prof. Mag. Dr. Thomas Bugnyar |
| Mitbetreut von \| Co-Supervisor: | Mag. Dr. Lisa Horn-Péter |

## Acknowledgements

First and foremost, I would like to express my sincere gratitude to my co-supervisor, Mag. Dr. Lisa Horn-Péter, for her patient guidance, clear and precise explanations, and exceptionally timely feedback throughout the entire process. I am also deeply indebted to Jozsef Arató, B.Sc., M.Sc., PhD, for his methodological advice on programming and machine learning. My appreciation further extends to Mag. Gabriela Marková, PhD, for her valuable insights into developmental psychology. I am very grateful to Remya Thottakara, M.Sc., for carrying out the data collection and generating the original Loopy coordinates. I would also like to thank the children who participated in the study, as well as their parents, without whose support this work would not have been possible. Finally, my personal thanks go to my friends and those close to me for their encouragement and support during this work.


**Disclosure statement regarding the use of generative artificial intelligence**

Generative AI (ChatGPT) was used for grammar revision and for support in code debugging. All outputs were verified by the author.

## Abstract

Since Hall's work on proxemics, researchers have argued that spatial closeness and bodily orientation toward one another are indicators of social relationships. Affiliation cues in early childhood are difficult to quantify because children's social behaviors unfold fast, involve multiple partners and coordinated movements. While wearable sensors and automated tracking have enabled detailed measures of spatial cues in naturalistic settings, few studies have applied predictive modeling to these data. Therefore, I investigated whether simple spatial measures, such as dyadic distance and social orientation angle, could be used to detect pre-schoolers' social interactions and tested the effect of friendship on these behaviors. I found that befriended dyads maintained significantly shorter distances than non-befriended dyads and engaged in longer social interactions, whereas differences in dyadic orientation angles were not statistically significant. Predictive logistic regression models confirmed these findings, indicating that shorter distances increased the probability of the social interaction class. The dyadic social orientation angle contributed additional explanatory power. Models including both predictors achieved the highest classification accuracy (~83%) and AUC (~0.9), however, models including distance as the sole predictor performed nearly as well. The convergence between inferential and predictive approaches highlighted proximity as a reliable cue of affiliation and demonstrated the potential of machine learning for automated detection of children's social interactions. Beyond advancing research methodology, this approach could drive early educational interventions by providing scalable and objective tools for monitoring social engagement in naturalistic group settings.

## Introduction

In many species, social interactions between group members are non-random and reflect social relationships characterized by their relative quality and distribution (Hinde, 1976). These patterns contribute to social complexity, which may represent the main selective force driving the evolution of large brains and intelligence among extant animals (Johnson-Ulrich, 2017). Within-group social relationships are heterogenous, fluctuate in strength and frequency over time, are influenced by ecological context, group composition, and individual state. They can serve different functions reflecting adaptive strategies (e.g., comparative review of group-living species, including primates, other mammals, birds and some invertebrates; Kutsukake, 2009). Affiliative relationships, in particular, contribute to group cohesion and cooperation (e.g., Dunbar, 2012), reproductive success (e.g., Silk, 2007), and survival (Archie et al., 2014) in primates, and may extend beyond pair bonds and kin bonds, as shown for example in rooks (Boucherie et al., 2016). Levels of affiliation seem to influence social learning (Schwab et al., 2008) and attention (Scheid et al., 2007) in corvids. Affiliation may be especially important for long-term cooperation, as primates tend to engage in cooperative acts with socially bonded partners (Sussman et al., 2005). This has for example been observed in wild chimpanzees, where food sharing was best predicted by enduring grooming relationships rather than coercion or direct signaling (Samuni et al., 2018). At the same time, elevated oxytocin levels after sharing likely reinforced cooperative bonds (Samuni et al., 2018). Grooming is a well-documented affiliative behavior which is often exchanged reciprocally in primates, along with agonistic support (Schino, 2007). Reconciliation mechanisms such as grooming coercion and the trading of social services have been identified as forms of relationship repair and post-conflict affiliation in wild macaques (McFarland & Majolo, 2011). The role of affiliation might also be implied in social support networks in the form of the long-term reciprocation of agonistic support in ravens (Fraser & Bugnyar, 2012), the formation of two-level-alliances in dolphins (Connor et al., 1999), and the use of serial aggression in social support-dependent rank in geese (Scheiber et al., 2009). Social bonds may further influence how individuals signal to one another. For example, vocal convergence, where individuals modify their calls to resemble those of their social partners, has been observed across taxa (e.g., Loth, 2007; Luef et al., 2017; Tyack, 2008). At the same time, ravens may suppress the usual food calls in the presence of an affiliation partner to prevent non-affiliated individuals from discovering the source (Szipl et al., 2015). Overall, affiliative relationships are an important component of complex social structure and manifest in specific affiliative behaviors.

The quality of affiliative relationships might be described by multiple dimensions which are not exclusive to one species. Fraser, Schino, et al. (2008) analyzed nine behavioral variables of captive adult chimpanzees (i.e., grooming, grooming symmetry, consistency of affiliation, proximity, tolerance to approaches, support, counter-intervention, aggression, and begging) using principal component analysis and identified three main components of relationship quality (i.e., value, compatibility, and security). These components had already been proposed by Cords and Aureli (2000) in the context of post-conflict reconciliation behavior. The *value* of a relationship is defined by direct benefits, such as agonistic support or food sharing, and it was later linked to consolation in chimpanzees (Fraser, Stahl, et al., 2008). *Compatibility* of relationship reflects the level of mutual tolerance and general tenor of interactions. The *security* dimension reflects the predictability and consistency of interactions over time. Interestingly, Fraser and Bugnyar (2010) extracted the same components via principal component analysis from seven behavioral variables in ravens (i.e., contact sit, preening, agonistic support, tolerance to approaches, aggression, counter-intervention, and variation in response to approach over time). The convergence of findings across primates and corvids suggests that dimensions of affiliative relationship quality may build on shared principles across diverse species.

While the assessment of affiliative relationships in non-human animals relies on behavioral cues (e.g., Fraser, Schino, et al., 2008), the research on affiliative relationships in human animals, particularly adults, is often conducted using verbal reports (e.g., six factors of ideal friendship expectations by Hall, 2012: enjoyment of time spent together, symmetrical reciprocity, instrumental aid, similarity, communion, and agency). However, the assessment of affiliative relationships in pre-school children using verbal reports (e.g., Selman, 1981) might lead to false negative results. Despite the association with relevant outcomes, such as academic achievement (Wentzel et al., 2018) and social and emotional well-being (Erdley & Day, 2017), research on affiliation in pre-school-aged children remains underdeveloped. Since the onset of independent locomotion plays crucial role in advancing spatial cognition and social communication (Campos et al., 2000), collecting spatiotemporal behavioral data, might result in more reliable findings for this age group.

In Chapter 10 of his book called *The hidden dimension*, Hall (1966) described four proxemic zones of social interaction. First, *intimate distance* (0 – 45 cm) with stepped up sensory inputs of olfaction, heat, sound, feel of breath and distorted sight that can be overwhelming (e.g., lovemaking, wrestling, crowded elevators in subway). Second, *personal distance* (45 – 120 cm) up to arm's length, from easy touching to a point when two people can touch fingers if they extend both arms. This zone enables discussion of mutual interests and mutual involvement, the vision is sharp, capturing face and movement of hands. Interestingly, it also corresponds to *personal distance* described in non-human animals by Hediger (1950) as distance which “non-contact species” maintain between themselves and their fellows. Third, *social distance* (120 – 360 cm), at which no touch is possible unless special effort. The details of the face are lost, however, eyes and mouth, skin texture, hair, condition of teeth and clothes are in the area of sharp vision. In this zone, impersonal business (e.g., reception at the hotel) or social gatherings take place. Hall observed that at this distance it is more important to maintain eye contact during conversations than at closer distances. Finally, *public distance* (from 360 cm) allows individuals to take evasive or defensive action if needed. While iris color is not perceivable, the white of the eye, which is indicative of gaze direction, is visible. Other persons can be seen peripherally. This zone is typical for public occasions where the voice of the speech must be amplified. Non-amplified communications shifts to gestures and body stance. In non-human animals, this zone correspond to *social distance* described by Hediger (1950) as distance, after exceeding of which, the ”non-contact species” no longer see, hear and smell their group and start feeling anxious. Hall (1966) remarked that this distance is increased in human animals by the use of media. *Personal space* is another, rather context-dependent construct, defined by Hayduk (1987) as the area at which intrusion of other people arouses discomfort. This intrusion is believed to cause withdrawal rather than a fight for dominion (Katz, 1937). Interestingly, Hediger (1950) saw positive correlation between size of the animal, its dominance (McBride, 1964) and fighting distance. Similarly, children seem to use more space as they grow older (Hayduk, 1983), acquiring adult distances by the age of twelve (Aiello & De Carlo Aiello, 1974). Aiello & De Carlo Aiello (1974) also emphasized the lack of sex differences until early adolescence. Similarly, Hayduk (1983) summarized previous findings regarding sex differences as inconclusive and confounded by setting, posture, or role expectations.

Recent research supported the sensitivity of pre-school children to social distance cues. For example, Paulus (2018) examined whether children at age 4 to 6 share similar expectations about interpersonal space as adults. He showed them sequences of images depicting two individuals standing apart, with one gradually approaching the other. At each step, participants judged whether the distance was still comfortable. In this study, pre-school children demonstrated clear expectations regarding interpersonal space, preferring smaller distances when the two individuals were friends and larger distances when they were strangers. Adults displayed the same pattern, though they generally preferred greater interpersonal distances overall. Moreover, Afshordi and Liberman (2021) inferred pre-school children's (3-6 years) understanding of social distance in the context of affiliative relationships. The participants were looking at pictures of three unknown children, and successfully identified who is friends with whom, based on proximity, prosocial interactions and similarity cues.

Another behavior linked to affiliative relationships might be eye contact. In the context of visual code, Hall (1963) described retinal combinations relevant in social contact: (1) sharp *foveal* vision (1°) due to 25,000 cones in fovea, each of which relay to a single bipolar nerve; (2) clear *macular* vision (3° vertically and 12° horizontally) due to presence of all cones and half of the rods at macula; (3) *peripheral* vision which allows to perceive motion laterally on temporal side of each eye; and (4) no visual contact, exceeding horizontal temporal vision of approximately 100° (Spector, 1990). Closely related to retinal combinations by Hall is the framework of joint attention presented by Siposova and Carpenter (2019). They summarized typology of joint social attention levels from literature to: (1) *monitoring attention* with one person attending to the observational act of the other person; (2) *common attention* of both persons looking the same thing; (3) *mutual attention*, when both persons are looking at the same thing and are reciprocally aware of each other's attention; and (4) *shared attention*, representing full common awareness of the other's knowledge about one's knowledge about each other's attention, allowing for communication about the shared subject of attention. In non-human animals, many species are able to follow gaze to locations outside their own visual field, suggesting joint attention at some level of understanding of others' ability to attend to things. For instance, Bugnyar et al. (2016) created cross-modal caching paradigm with peephole and playback of a raven conspecific, requiring transfer of own prior visual access to unseen competitors.

While eye contact is potent, body orientation is easier to measure and could serve as proxy for eye contact (Mehrabian, 1968). Osmond (1975) described sociofugal body orientation that decreases the probability of social interactions, and sociopetal orientation that increases probability of social interaction. Building on this concept, Hall (1963) proposed a notation system for recording and comparing social orientations. In his sociofugal-sociopetal-axis, he differentiated several body orientations, such as: (1) face-to-face orientation provides maximal sociopetality, allowing for direct communication reaching each other with maximal intensity; (2) perpendicular facing orientation which is more casual and less involved, allowing to discuss subject of common interest; (3) parallel shoulder-to-shoulder orientation, allowing informal transitory communication or simultaneous watching; and (4) back-to-back as most sociofugal orientation. He emphasized that proxemic communications are always embedded in context (i.e., culture, social setting, age, status and sex) and must be read as such. Kendon (1977) defined an *F-formation* in which participants' lower bodies are turned towards a common center, to which they have equal, direct and exclusive access.

Evidence showed that multiple proxemic behaviors act together. For example, McBride et al. (1965) measured galvanic skin response from participants who were approached by experimenter. They measured highest response in persons who were approached frontally, intermediate response from participants approached from the side, and least response from participants approached from the back, confirming that greater arousal levels are associated with approaching more direct body orientations. In the same year, Argyle and Dean (1965) introduced the equilibrium theory focused on balancing social intimacy to tolerable level by reducing eye contact in closer proximity. Also, Fischer (1968) asked adults to position human figures on a board and found that figures facing each other were placed farther apart than those facing away from each other or in parallel. Fischer's results were later replicated in children (Estes and Rush, 1971). Cristani et al. (2011) provided the first statistical implementation of Kendon's *F-formation*, showing that groups can be detected from spatial orientation cues alone. Few years later, Setti et al. (2015) introduced a graph-cuts optimization method and demonstrated that positions and body orientations are sufficient to detect free-standing conversational groups in images. Also, Lahnakoski et al. (2020) found that increased distance between participants predicted lower enjoyment, while increased facing orientation correlated with cooperative effort reported by participants, also supporting the importance of proxemic cues of social distance and social orientation angle combined. The authors even reported that facial and bodily orientation were more predictive of subjective evaluations of the interaction than specific gaze behaviors.

Proxemic behaviors in young children are particularly challenging to quantify during natural interactions (Horn et al., 2024). Observational and video-based methods often cannot follow the interactions of several children at once (Banarjee et al., 2023), creating a trade-off between capturing detailed behavioral information for a small number of children or collecting more basic measures from a larger group. In addition, traditional observation is constrained by the subjectivity of human coders, time required for continuous behavioral coding, and the difficulty of registering short-lived or subtle interaction cues in fast-moving group settings.

One approach for overcoming these limitations is the use of lightweight wearable devices or automated video tracking, allowing for precise spatiotemporal recording of whole groups (Mukhopadhyay, 2014). For example, Dai et al. (2020) utilized badges that exchanged low-power radio signals to log face-to-face proximity events and tracked hundreds of children (aged 2.5-6.5 years) during free play in a pre-school context without restrictions. Messinger et al. (2019) employed radio-frequency identification (RFID) to record location of 5-year-old children in kindergarten classrooms. Then, using radial distribution function, they compared the actual dyadic distance to dyadic distance expected by chance, and identified social contacts occurring at above chance level. Similarly, Drye et al. (2024) also used RFID tags to collect over 750 hours of high-precision (ultra-wide band) location and orientation data from pre-schoolers (3-5 years) and their teachers. They identified differences in social approach velocity and amount of time spent in social contact between children with and without autism disorder. Eichengreen et al. (2024) used lightweight chest-mounted proximity badges to examine schoolyard interactions between deaf and hard-of-hearing preadolescents and their hearing peers (both aged 8-14 years), detecting moments of co-location and mutual orientation. Other approaches focused on dyadic coordination. For example, Li et al. (2023) detected joint attention and mutual gaze from videos of free play between children (approximately 10 months old) and their parents by tracking their head positions, detecting objects and generating heatmaps of estimated gaze directions, thus avoiding the use of unwieldy eye-tracking hardware. Similar methodical arrangements also be applied for adult sample. For example, Lahnakoski et al (2020) captured 3D positions of forehead and dominant hand on two depth cameras to compute social orientation, dyadic proximity, synchrony and gaze behaviors (approximated from head orientation). They found that dyadic proximity and orientation towards each other were linked to enjoyment of a joint task and cooperation effort. This finding aligns with previous research showing that affiliated peers engage in longer social interactions (e.g., Fabes et al., 2003; Hartup et al., 1988; Howes, 1983; Newcomb & Bagwell, 1995).

While these studies employed automated tracking, none of them explored the next logical step of predictive modelling based on measures already derived from the data. Horn et al. (2024) highlight the advantages of combining wearable sensor technology with machine learning to classify and quantify complex social movement patterns embedded in space, time and different social contexts. Machine learning has the potential to move behavioral research to a new level by analyzing large, complex datasets collected from novel sources, and detection of patterns that might be non-linear or involve multiple interacting variables (Turgeon & Lanovaz, 2020). Moreover, by shifting the focus from explaining behavior to predictions, machine learning can help behavioral science develop models that generalize better to future real-world behaviors (Yarkoni & Westfall, 2017). Some researchers already attempted to classify behaviors in natural settings (e.g., walking vs. sitting, light vs vigorous activity) using random forests (e.g., Ahmadi et al., 2020) or artificial neural networks (e.g., Hagenbuchner et al., 2015). Relevantly, Dai et al. (2020) applied sequence learning model (bidirectional long short-term memory network) to sequences of dyadic proximity cues recorded by RFID sensors to predict social interaction class using temporal patterns. They found that proximity combined with face-to-face positioning improved the probability of correct classification of social interactions. Taken together, previous research suggested that predictive modelling based on spatial and temporal measures is a promising next step for advancing automated analyses of social interactions.

The aim of the present work was to evaluate whether machine learning models can be applied to spatial measures of dyadic proximity and social orientation to predict social interactions in naturalistic kindergarten setting. I extracted measures of proximity and head orientation from automated tracking data of a free-play, combined these data with friendship assessments for additional analyses, and validated the automated measures against hand-coded observations of social interactions. I hypothesized that befriended dyads would display higher spatial proximity and head direction closer to face-to-face orientation compared to non-befriended dyads. To further establish the link between friendship assessments and social interactions predicted by the algorithm, I examined the effect of friendship on the duration of social interactions. Specifically, I expected longer social interaction durations in befriended dyads compared to non-befriended dyads. Finally, I explored whether machine learning algorithms could successfully detect hand-coded social interactions in dyads from the spatial measures of proximity and orientation.

## Method

In this project, I worked with existing video and data material collected by Thottakara (2022) between September 2021 and April 2022 as part of her master's thesis. In her study, she investigated interpersonal synchrony in pre-school children during free play, focusing on how friendship influenced behavior. For this purpose, she recruited participants, videorecorded their behavior in a kindergarten facility, assessed their friendships, and obtained tracking data of their positions in the room using the commercial tracking software *Loopy* (loopbio GmbH, Vienna, Austria; https://loopbio.com/loopy/). I extended this material by annotating the videos for social interaction behaviors, correcting for spatial distortion in the tracking data, and computing new spatial measures (i.e., dyadic social orientation and dyadic proximity). Then I investigated the effect of friendship on the spatial measures, and on the durations of social interactions. Finally, I applied machine learning algorithms to explore how well the spatial measures could predict the occurrence of social interactions among pre-school children.

**Participants**

A total of 20 pre-school children (10 female, 10 male), aged between 3 and 6 years (*M* = 3.95 years, *SD* = 0.82), had been recruited at a company childcare facility associated with a research institution and a media hub in Vienna, Austria. All participants came from an urban environment and had international backgrounds, most originating from Europe and some from non-European countries (e.g., USA, India). While no data was gathered on parental education, socio-economic status (SES), or the children's ethnicity, the association with a research institution and a media hub suggests that families were likely from middle to high SES and educational backgrounds.

All procedures were carried out in accordance with the Declaration of Helsinki and had been approved by the ethics committee of the University of Vienna (Ref. No. 00599). Prior to the study, informed consent was obtained from all participants' parents or legal guardians and the children's participation in the study was voluntary.

**Friendship assessment**

All participants were tested separately in familiar rooms of the childcare facility. They were already familiar with each other and passed a photo-name recognition test. They were administered a modified version of the Peer Preference Assessment (Birch & Billmann, 1986) to determine degree of liking among each other. First, each child was asked to assign peer's photograph to cartoon faces representing three categories: (a) happy – "I enjoy playing with this child and I often play with this child," (b) sad – "I don't enjoy playing with this child and I rarely play with this child," and (c) neutral – "Playing with this child is okay, I don't really care." As a second measure, children were asked to position an arrow on a sliding scale with a sad cartoon face on the left and a happy cartoon face on the right side, to indicate how much they enjoyed playing with each peer. The scale was then divided into three equal sections to assess the degree of liking: (a) near the happy face, (b) near the sad face, and (c) the middle section. In both categorization measures, ratings in section (a) were interpreted as indicative of friendship, while ratings in section (b) were considered to reflect peer dislike and were labeled as non-friend.

Participants were first grouped into dyads based on mutual ratings across both measurements. Twelve children mutually rated each other as a friend on both scales, resulting in 12 focus dyads (4 female-female, 3 male-male, and 5 mixed-sex). For each of these dyads, one additional child was identified who was rated as a friend by both dyad members, resulting in a so-called *F-constellation* (a triad consisting of 3 dyads: befriended, befriended, befriended). Also, for eleven of the focus dyads, one additional child was identified who was rated as a non-friend on both scales (*N* = 7), or in three out of four total ratings (i.e., each child gave one rating on two scales, *N* = 4), forming a so-called *NF-constellation* (a triad consisting of 3 dyads: befriended, non-befriended, non-befriended). Accordingly, each focus dyad was supposed to be tested twice, once in the *F*-constellation and once in the *NF*-constellation. However, one focus dyad was not tested in *NF*-constellation for organizational reasons. The formation of both constellation types by adding a child to a focus dyad is illustrated in Figure 1. The sex composition of the resulting groups was 4 all-female triads, 2 all-male triads and 17 mixed sex triads. Most children participated in multiple triads (*Mdn* = 2 triads, range = 1–10). The order of observations for each constellation was randomized.

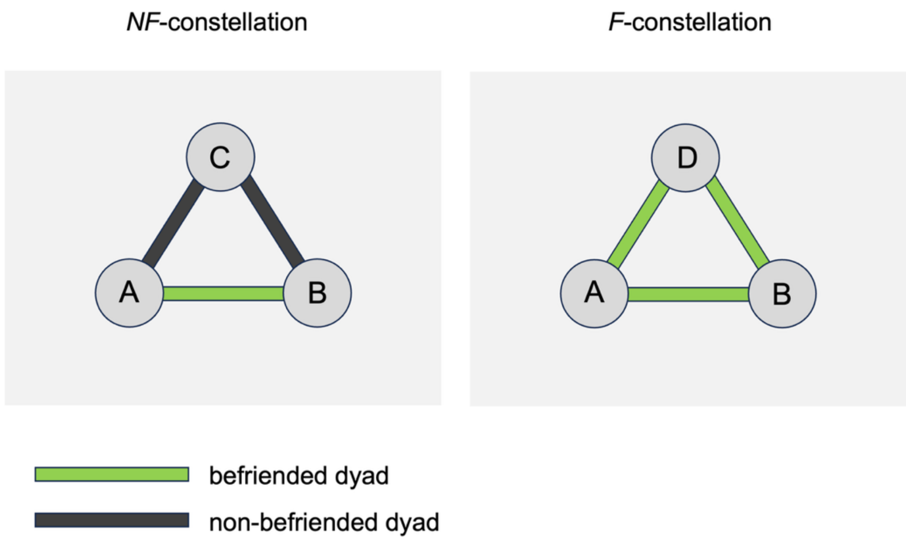



**Figure 1. Two types of groups were created based on friendship assessment.** The *NF*-constellation groups were formed from focus dyad (A and B) by adding mutually disliked peer (C), whereas the *F*-constellation groups were created by adding mutually liked peer (D). Befriended dyads are depicted by a green relationship line, while non-befriended dyads are represented by black line.

**Video material and dataset with positional tracking data**

Each triad was recorded on video during the ten minutes of a free play session in a room equipped with age-appropriate toys. The room size varied approximately from 2 x 3.5 m for the smallest room (observation DYAD02F) to 4 x 6 m for the biggest room (observations DYAD06F and DYAD06NF). The overhead camera GoPro HERO 8 Black (v2.50; 60 fps) was placed on the ceiling in the middle of the room. The experimenter took a seat in a corner of the room and deliberately avoided involvement in children's play. Children were wearing colored caps with two markers (i.e., *front, center*) and the opening in the back of the cap (i.e., *back*). These markers were observable on the recording made by the overhead camera from above. After obtaining the video recordings, they were first uploaded to *Loopy* (loopbio GmbH, Vienna, Austria; https://loopbio.com/loopy/) and corrected for radial lens distortion using a chessboard calibration method. The positions of the markers on the 2D-plane were then manually annotated in *Loopy* by Remya Thottakara in 800 still frames across 10 videos. Built-in deep learning algorithm in Loopy further predicted for remaining frames using the commercial tracking software *Loopy.* This procedure resulted in a dataset of XY-coordinates defined by the cap markers worn by each child, thereby indicating positions of their heads.

**Correction for peripheral distortion**

Despite the intrinsic chessboard calibration applied in Loopy, the video recordings were distorted in a way that led to seemingly longer distances between participants, particularly above the floor level at the periphery. This was especially problematic for automated tracking data, where markers on children's caps were detected by an algorithm. To correct for this distortion, a 2D linear transformation function was calculated and applied to all Loopy coordinates. The room depth dimension was approximated from a table which was present in each recording. The XY-coordinates of three tabletop corners (i.e., approximately at the height of sitting children) and at the base of three visible table legs were hand-selected for each video separately using a mouse callback function implemented in *OpenCV* (version 4.10.0; Bradski, 2000) in Python (version 3.13.0; Python Software Foundation, 2023). The coordinates were translated to bottom and top overlay grid, respectively (Figure 2). The grids were aligned by the center of the frame and the vectorial difference between the grid nodes was fit to a linear regression model using *Scikit-learn* (version 1.6.0; Pedregosa et al., 2011) and modelled as a radial pull distortion in the Cartesian coordinate system (Figure 3; Appendix A). Figure 4 illustrates the shift of the right-most datapoint of the video DYAD10NF. Data organization and tabular output were supported by *pandas* (version 2.2.3; The pandas development team, 2020; McKinney, 2010) and *tabulate* (version 0.9.0; tabulate, 2024), *NumPy* (version 2.1.3; Harris et al., 2020) was used for numerical operations and *Matplotlib* (version 3.9.2; Hunter, 2007) for visualization.

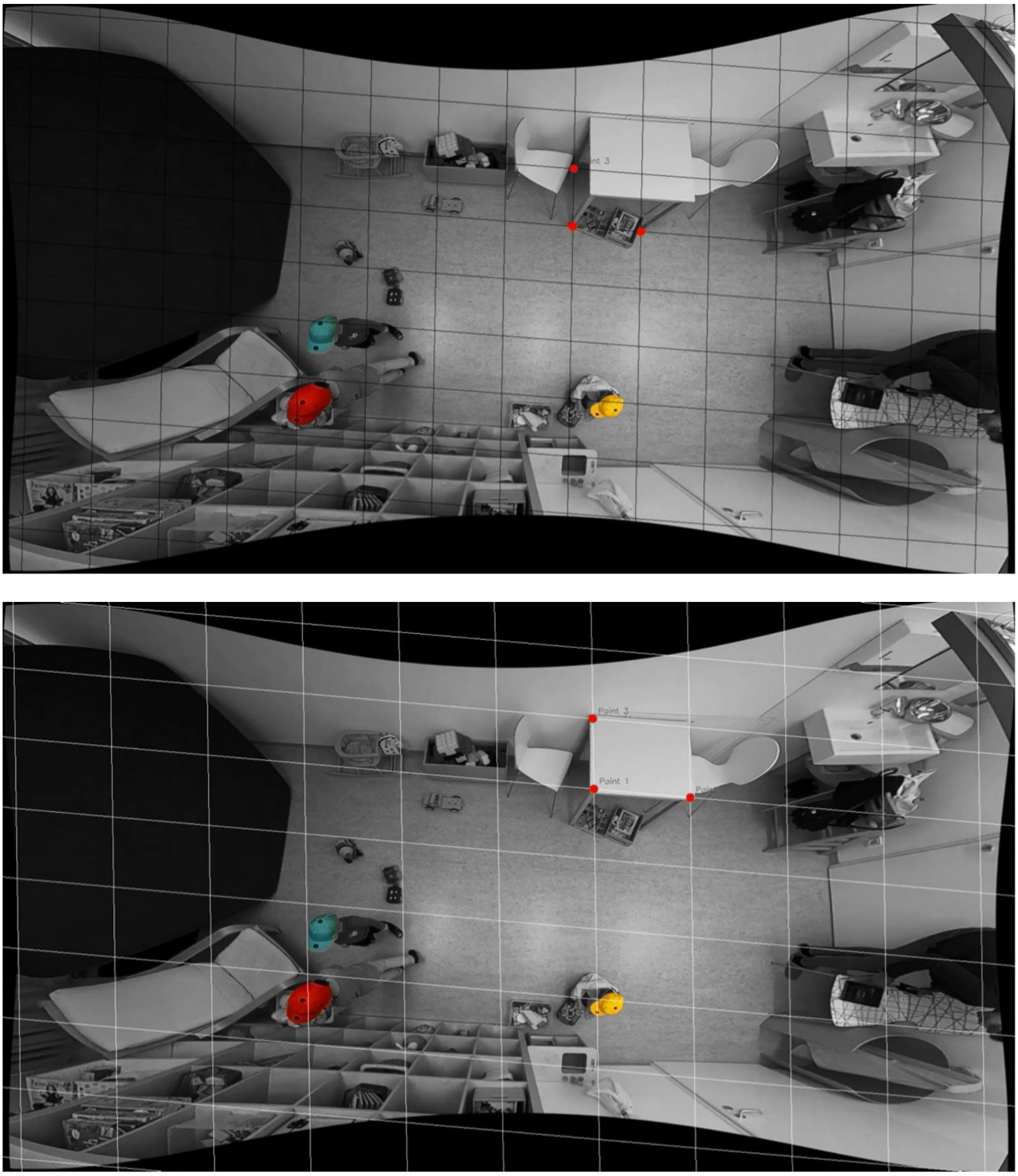



**Figure 2. Bottom and top grid derived from hand-selected points.** An example from the first frame of the video DYAD10NF. A: Hand-selected points at the base of table legs (red) serve as nodes of the bottom overlay grid (black). B: Hand-selected points at the tabletop (red) serve es nodes of the top overlay grid (white).

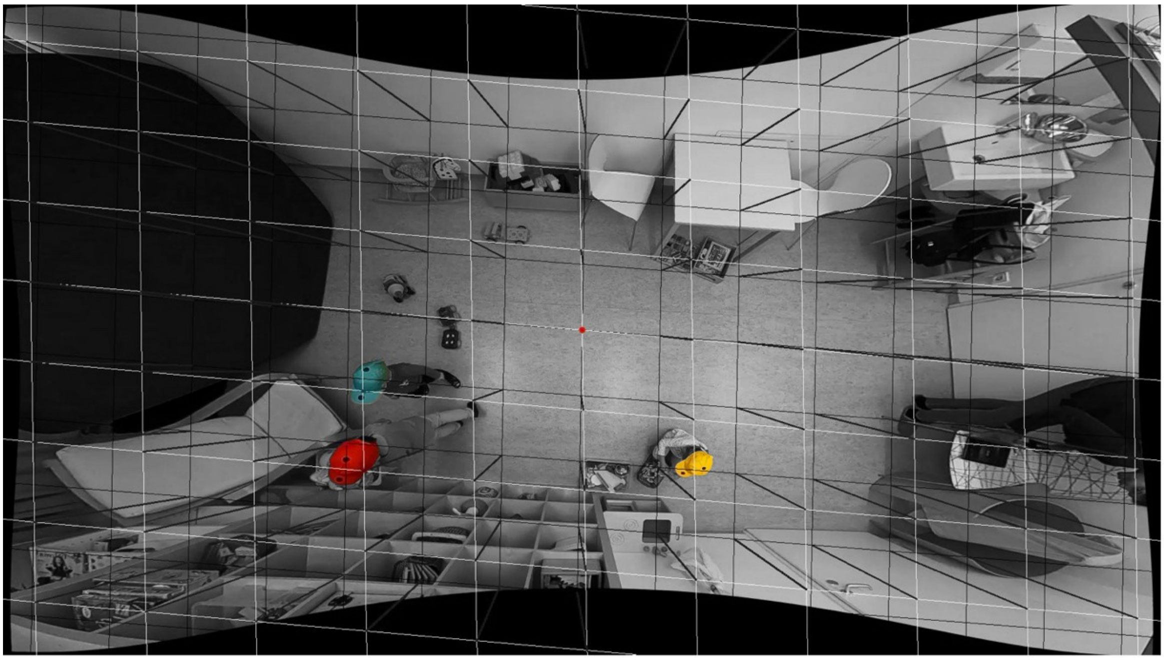


**Figure 3. Transformation computed from vectorial grid nodes difference.** An example from the first frame of the video DYAD10NF. The top and bottom grid were aligned by the center of the frame and vectorial distance between corresponding nodes has been computed (depicted as radial thick lines connecting grid nodes, red in the middle).

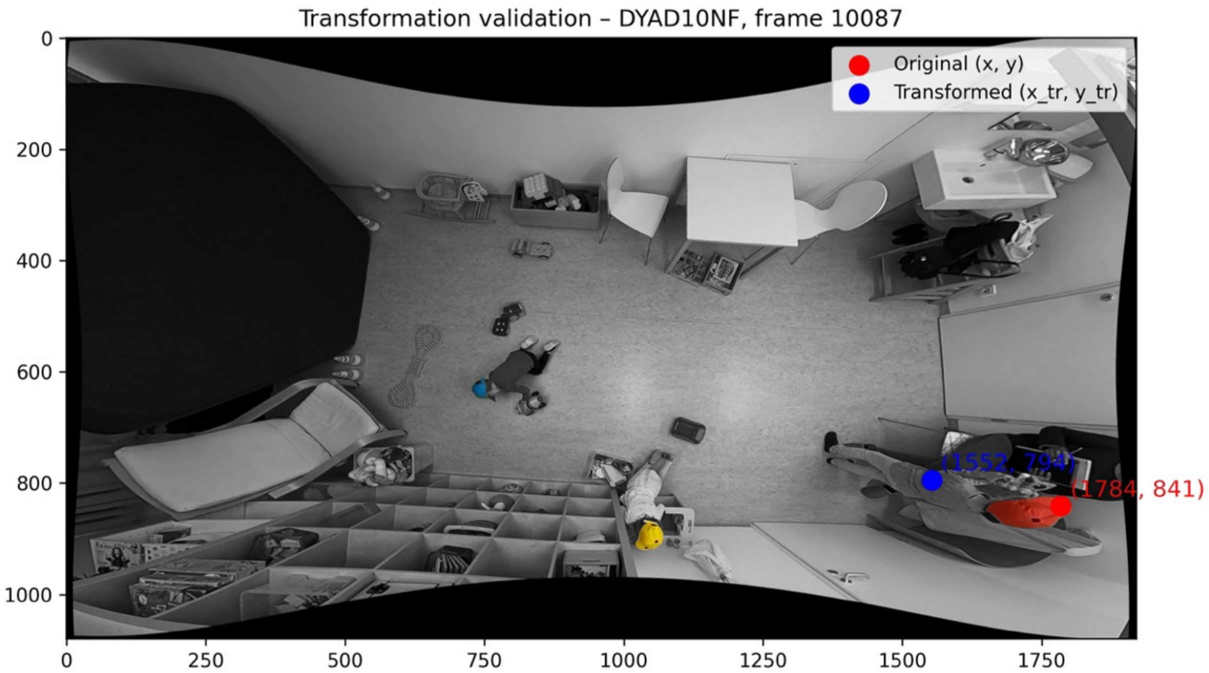



**Figure 4. Visual validation check for the direction of the radial pull.** In this example, the right-most datapoint (front marker at the red colored cap) located at $x$ = 1784.20, $y$ = 841.12 (red) was transformed to $x_t$ = 1552.43, $y_t$ = 794.28 (blue). This transformation represents a radial pull at the level of the top grid node in the direction to the level of the bottom grid node. The coordinates were overlaid over the corresponding frame number 10087 with the right-most datapoint in the video DYAD10NF. The x-axis represents the x-coordinates of the standard top-left-origin (non-Cartesian) screen space, and the y-axis represents the corresponding y-coordinates in that same system.

### Creating spatial measures from XY-coordinates

To address missing positional data on 100 ms intervals, I downsampled the spatial coordinates by averaging them over 250 ms intervals. Then I interpolated values that were still missing using SciPy's *interp1d* function (Virtanen et al., 2020). For the missing values at the beginning and at the end of the dataset I applied extrapolation within the same function. I excluded datapoints referring to the opening at the back of the cap (*back* marker) due to redundance and disproportionally many missing values. Finally, I cut all recordings to 10-minutes-length to ensure uniform length across different observations.

One construct relevant for the study was the dyadic proximity. I operationalized it as Euclidean distance between the averaged coordinates of *front* and *center* markers of two participants (Figure 5). To approximate real distances in millimeters, I scaled distances for each recording using a scaling factor (px to mm) based on the conversion of pixels to the length of a visible tabletop in centimeters.

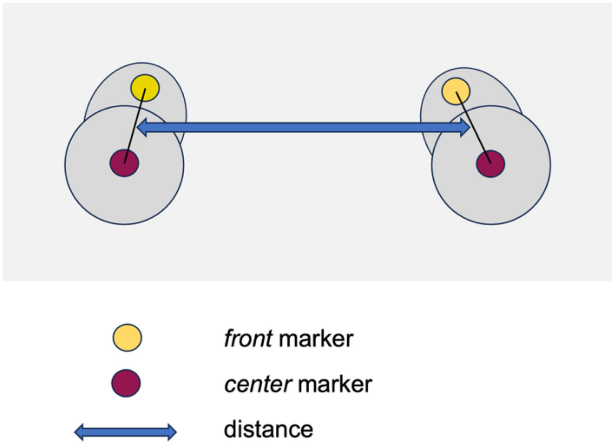
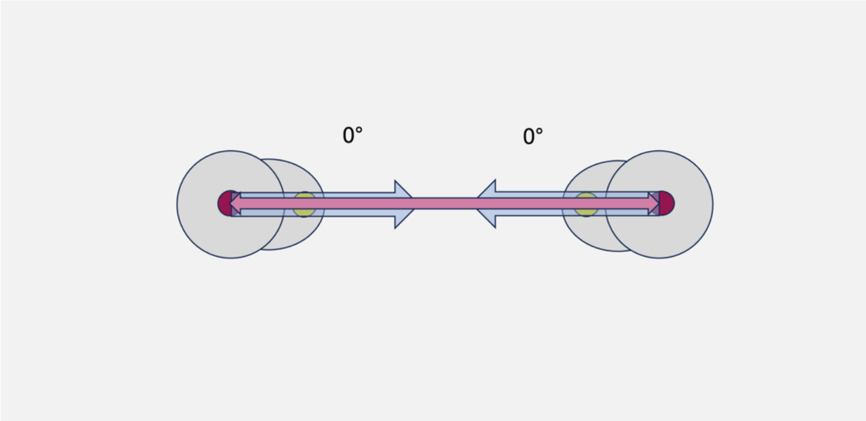



**Figure 5. Dyadic distance computed from relative difference in averaged cap markers' position***.*

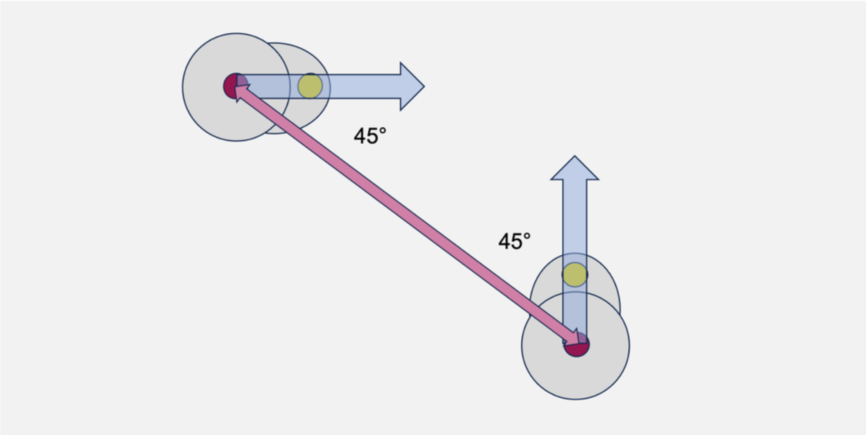
Schematic illustration of dyadic distance measure (blue arrow) as a difference between averaged coordinate position of *front* (yellow dot) and *center* (purple dot) markers of a dyad.

Another construct relevant for the study was dyadic social orientation. Similar to the interpersonal face orientation measure described by Lahnakoski et al. (2020), I operationalized it as an angle reflecting the degree to which two children deviated from facing each other. To compute this angle, I used the XY-coordinates of the *center* and *front* markers of the dyad. First, I defined a baseline vector pointing from one child's *center* marker to their partner's *center* marker. I then calculated the angle between this baseline vector and the orientation vector extending from one child's *center* to their *front* marker, indicating their head orientation. Specifically, I divided the dot product of the two unit-vectors by the product of their magnitudes, I clamped the cosine value between −1 and 1 to avoid rounding errors, and then I converted the result from radians to degrees. Next, I calculated the angle between the baseline vector and the orientation vector of the second child using the same pipeline. This way, I computed two separate angles, one for each child, which I summed to generate a dyadic social orientation ranging from 0 to 360 degrees, with low values indicating social orientation towards each other (Figure 6).

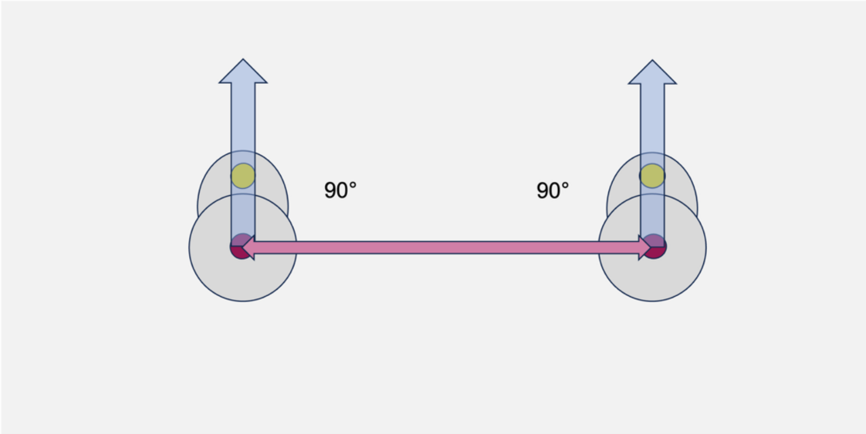
(A) Facing each other completely corresponds to dyadic social orientation of 0°




(B) Facing each other perpendiculary corresponds to dyadic social orientation of 90°




(C) Parallel orientation in the same direction corresponds to dyadic social orientation of 180°



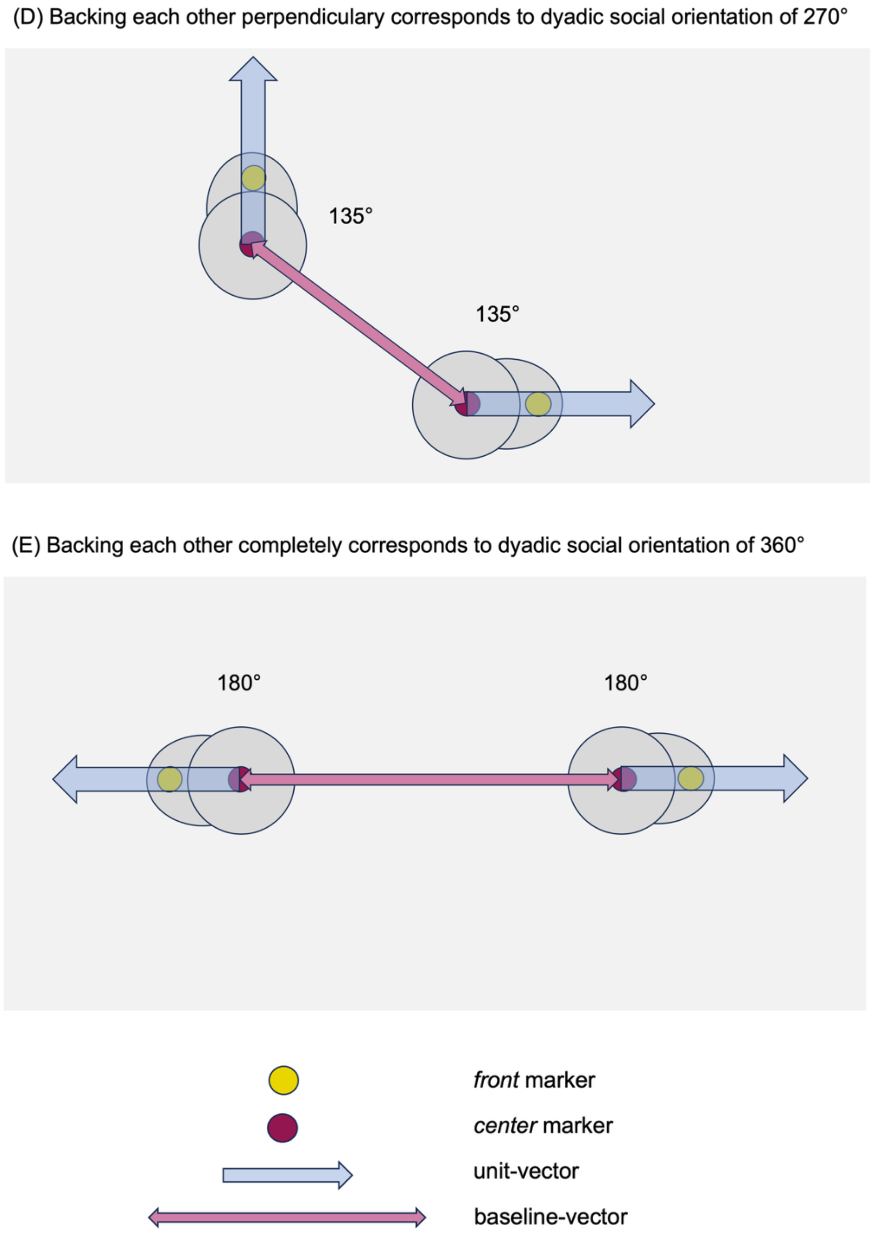



**Figure 6. Social orientation angle.** Dyadic social orientation was defined by the angle formed by baseline vector (pink arrow) and two orientation vectors (light blue arrow). Orientation vector of each child is starting from the *center* marker (purple dot) and following the direction to the *front* marker (yellow dot) of the cap.

**Behavioral coding**

I used Solomon Coder (version beta 19.08.02; ©András Péter) to manually annotate the instances of social interaction behavior at a rate of 100 ms. I was unaware of the relationship assessment between children. I divided social interactions to *clear*, *distant* and *unclear* social interaction categories. I coded *clear* social interactions whenever children communicated, played and/or fought under the following two preconditions: (1) reaching proximity (i.e. at arm’s length, extending up to the far phase of personal distance, as described by Hall, 1966), and (2) body and/or head orientation towards one another, which served as a proxy for eye contact (i.e., 1°, within sharp vision of the foveal retinal combination, as described by Hall, 1963). This category was the least ambiguous and was used as a most conservative measure of interaction. Because children often engaged and disengaged from interaction, instances that disappeared and reappeared within a 10-second interval were coded as continuous. To account for more ambiguous signals and increase sensitivity, I introduced a complementary category of *unclear* social interaction. I coded *unclear* interactions if both preconditions were met, but it was unclear whether interaction truly took place, or when children showed behavior which I considered social interaction despite violating the precondition of orientation. However, one-sided behaviors, such as following, interacting in parallel with a third party, sitting butt to butt, jointly observing or pointing on the same object, were not considered social interaction. The category of *unclear* social interaction was conceptually linked to *clear* interactions, not meaningful on its own, and was therefore not evaluated or analyzed separately.

Since social interactions can also occur over distance (e.g., Hall, 1966), I also introduced the category of *distant* interactions. I coded *distant* interactions, when the precondition of proximity was violated but social orientation was present, such as throwing objects to each other or talking over a distance. However, due to their scarce frequency, *distant* interactions were not analyzed as a separate outcome variable. Instead, I combined all three interaction types (*clear, unclear* and *distant*) to create a global measure of interaction.

An inter-rater reliability test for durations within each minute of six randomly chosen videos (26% of all videos, four videos from *NF*-constellation and two videos from *F*-constellation) revealed an excellent agreement for the category of *clear* social interaction (ICC = 0.791, $F(179,122) = 9.06$, $p \leq 0.001$), even better agreement for the category combining *clear* and *unclear* social interaction (ICC = 0.828, $F(179,179) = 10.60$, $p \leq 0.001$), and good agreement for all *clear*, *unclear* and *distant* interactions combined (ICC = 0.707, $F(179,157) = 5.99$, $p \leq 0.001$).

**Data analysis**

To address the first part of the first hypothesis, I compared dyadic distances between befriended and non-befriended dyads. Since the *F*-constellation consisted only of befriended dyads, I limited the analysis to the *NF*-constellation, which contained both types of dyads. Each dyad contributed 2,400 datapoints across 10-minutes of free-play in one video observation session. Since I was not interested in temporal trends, I accounted for the within-dyad dependency by averaging the distance values across time. This resulted in one average distance value for one befriended dyad and two average values for the non-befriended dyads in each video observation session. I fitted the linear mixed-effects model using the *lmer()* function from the *lme4* package (version 1.1.33; Bates et al., 2015) in R. To account for the nested structure of the data, I included the video observation number as a categorical random factor. I compared model with random slopes against a model with random intercepts using Akaike's Information Criterion (AIC; Akaike, 1974). The model with random intercepts provided a better fit (AIC = 435.28) than the model with random slopes (AIC = 435.98), therefore I proceeded with the model without random slopes. Intra-class correlation coefficients (ICCs) of 0.78 and 0.89 suggested that including observation as a random factor accounted for a meaningful portion the total variance in both models. I tested the assumptions of the random intercept model using simulation-based residual diagnostics implemented in *DHARMa* package in R (version 0.4.7; Hartig, 2022). The residual diagnostics showed no evidence of overdispersion, non-normality, or heteroscedasticity (Appendix B, Figure B1). A separate Levene's test in the *car* package in R (version 3.1.2; Fox & Weisberg, 2019) also indicated no heteroscedasticity $F(1, 31) = 0.04$, $p = .847$. A residuals vs. fitted values plot (Appendix B, Figure B2) showed no fan-shaped pattern, further confirming that assumptions of linearity and homoscedasticity were met. Histogram of dyadic distances for befriended and non-befriended dyads is plotted in Appendix B, Figure B3. Influence diagnostics was computed using the *influence()* function with Cook's distance implemented in the *lme4*package in R.

To address the second part of the first hypothesis, I compared dyadic social orientation angles between befriended and non-befriended dyads. I again restricted the analysis to *NF*-constellation, since the *F*-constellation did not contain any non-befriended dyads. To account for the within-dyad dependency across the 2,400 datapoints during the 10-minute-long recording, I averaged the angles for each dyad, again resulting in three values for each observation session (of one befriended dyad and two non-befriended dyads. I fitted the linear mixed-effects model using the *lmer()* function from the *lme4* package. I included the video observation as a context factor. The AIC score was lower in model with random intercepts (AIC = 266.52, ICC = 0.48) compared to model with random slopes (AIC = 269.91, ICC = 0.52), suggesting that adding random slopes decreased the model fit. Therefore, I proceeded with the model including random intercepts only. The simulation-based *DHARMa* residual diagnostics showed no evidence of overdispersion, non-normality, or heteroscedasticity (Appendix B,

Figure B4). The additional Levene's test implemented in the *car* package was not significant either $F(1, 31) = 0.265$, $p = .610$. The residuals vs. fitted values plot (Appendix B, Figure B5) showed only slight curvature, further indicating the assumptions of linearity or homoscedasticity were met. Histogram of dyadic social orientation angles for befriended and non-befriended dyads is plotted in Appendix B, Figure B6.

To test the second hypothesis, I compared durations of social interactions between befriended and non-befriended dyads. Since the *F*-constellation consisted only of befriended dyads, I restricted the analysis to the *NF*-constellation. I a fitted linear mixed-effects model using the *lmer()* function from the *lme4* package, with observation session as random factor. To account for different aspects of social interaction separately and to address different ambiguity levels, I fitted three models with different types of social interaction as outcome variable: (1) *clear* interactions separately, (2) *clear* and *unclear* interactions combined, and (3) all interaction types combined.

In the models focusing on duration of *clear* interactions, the AIC score decreased (AIC = 341.20) and ICC score improved dramatically (ICC = 0.99) after adding random slopes to the model with random intercepts (AIC = 367.74, ICC = 0.36). Suggesting better model fit, with slopes variation across observation sessions accounted for large variance in the outcome. Therefore, I decided to use the random slopes model and tested its assumptions. Residual diagnostics using the *DHARMa* package indicated significant deviations from uniformity, but no overdispersion or heteroscedasticity (Appendix B, Figure B7, Panel A). The additional Levene's test from *car* package did not confirm heteroscedasticity either ($F(1, 31) = 1.86$, $p = .182$). However, the visual inspection of residuals vs. fitted values plot revealed a strong fan-shaped pattern (Appendix B, Figure B8, Panel A), indicating that the assumption of equal variance of residuals was violated. Since the distribution of *clear* social interactions was right-skewed (Appendix B, Figure B9, Panel A), I applied a log-transformation (Appendix B, Figure B9, Panel B) and re-evaluated the model assumptions. Fitting the model with log-transformed outcome and random slopes resulted in a singularity fit, also confirmed by very low intra-class correlation coefficient (ICC = 0.01), compared to model with random intercepts (ICC = 0.20). Since random slopes of observation contributed very little to variance of the outcome, I proceeded with the random intercept model despite slightly worse model fit (AIC = 121.08 vs. 124.56). The *DHARMa* diagnostics of this model was non-significant, raising no evidence of overdispersion, non-normality, or heteroscedasticity (Appendix B, Figure B7, Panel B). The separate Levene's test in the *car* package was non-significant ($F(1, 31) = 3.18$, $p = .084$) and the residuals vs. fitted values plot (Appendix B,

Figure B8, Panel B) showed no fan-shaped pattern, further confirming that the assumption of homoscedasticity was met. However, the LOESS line indicated a U-shaped pattern, possibly violating the linearity assumption. Since the assumptions for the linear mixed model were not fully met, I conducted nonparametric Wilcoxon signed-ranked test as an alternative. For this purpose, I averaged the duration of the two non-befriended dyads within each video observation, resulting in with one pair of comparison (befriended vs. non-befriended) per observation.

After including *unclear* interactions to the outcome in the linear mixed model, I compared the intraclass correlation coefficients and model fit of models with and without random slopes. The model with random slopes showed better fit (AIC = 368.23 vs. 379.84) and explained much more variance in between observation sessions (ICC = 0.948 vs. 0.343), so I decided to proceed with this model. *DHARMa* diagnostics suggested no violations of assumptions of normal dispersion, normality of residuals' distribution or homoscedasticity (Appendix B, Figure B10, Panel A). The additional Levene's test from *car* package was non-significant as well ($F(1, 31) = 1.39$, $p = .248$). However, the visual inspection of plot of residuals vs. fitted values indicated clear heteroscedasticity (Appendix B, Figure B11). Since the distribution of the outcome was right-skewed (Appendix B, Figure B12, Panel A), I applied a log-transformation to the outcome (Appendix B, Figure B12, Panel B) and re-evaluated the model assumptions. Fitting the model with log-transformed outcome and random slopes resulted in a singular fit, also observable was very low intra-class correlation coefficient (ICC = 0.02), compared to model with random intercepts (ICC = 0.21). Therefore, I proceeded with the random intercept model despite worse model fit (AIC = 122.71 vs. 127.62). *DHARMa* diagnostics (Appendix B, Figure B10, Panel B) and the additional Levene's test ($F(1, 31) = 9.20$, $p = .005$), both suggested significant heteroscedasticity, despite slightly improved distribution of residuals observable also in the residuals vs. fitted values plot (Appendix B, Figure B11, Panel B). However, the LOESS line indicated an U-shaped pattern, possibly violating the linearity assumption. Since the assumption of homoscedasticity for the linear mixed model was violated after log-transformation of the outcome, I conducted nonparametric Wilcoxon signed-ranked test as more robust analysis. Again, I averaged the duration of the two non-befriended dyads within each video observation, to allow for pairwise comparisons (befriended vs. non-befriended) per observation.

Last but not least, I analyzed the influence of friendship on durations of all types of social interactions combined, while accounting for the random variability between observations. Linear mixed model with random slopes showed better model fit (AIC = 374.94) than model with random intercepts only (AIC = 388.13). The comparison of ICC scores also suggested that random slopes accounted for a large portion of outcome variance (ICC = 0.95 vs. 0.25). I tested the model with random slopes for assumptions. *DHARMa* diagnostics revealed no evidence for heteroscedasticity, non-uniformity or over-/underdispersion of residuals (Appendix B, Figure 13, Panel A). The additional Levene's test from *car* package was non-significant ($F(1, 31) = 0.60$, $p = .443$). However, after visual inspection of the plot of residuals vs. fitted values (Appendix B, Figure 14, Panel A), which showed fan-shaped pattern, I decided to log-transform the outcome, which was initially right-skewed (Appendix B, Figure 15, Panel A & B). Fitting the model with log-transformed outcome and random slopes resulted in a singular fit (ICC = 0.00), that is why I proceeded with model with random intercepts only (ICC = 0.24), despite worse model fit (AIC = 126.57 vs. 116.86). Fitting the log-transformed model again led to non-linearity of residuals indicated by the LOESS line in the plot of residuals vs. fitted values (Appendix B, Figure B14, Panel B). Moreover, *DHARMa* diagnostics (Appendix B, Figure B13, Panel B) and the additional Levene's test ($F(1, 31) = 8.92$, $p = .005$), both suggested significant heteroscedasticity. Because of the violations of the assumptions for mixed linear model, I performed and reported the nonparametric Wilcoxon signed-ranked test over the pairs of befriended vs. (averaged) non-befriended dyads for each observation.

To explore whether machine learning algorithms can detect hand-coded social interactions based on relative distances and social orientation angles, I fitted mixed logistic regression models in R using the *glmer()* function from the *lme4* package and evaluated using a leave-one-out cross-validation (LOOCV) approach (Hastie et al., 2009), where I held each observation out from the training set once to use it as a test set. To account for the variance between recordings, I included the observation session as random contextual variable. Since there was no reason to exclude the constellations consisting only of befriended dyads from the model, I used data of both constellation types (*F*-constellation vs. *NF*-constellation) and included the constellation type as a fixed factor to the models. I compared performances of four models: (1) a model with distance as the only predictor, (2) a model with angle as the only predictor, (3) a model with both distance and angle as additive predictors without the interaction term, and (4) a model with both predictors including their interaction term. Since models including random slopes performed similar or better, models relying solely on random intercepts were not reported. To test performance for multiple definitions of social interaction, I repeated the whole procedure independently for the three different levels of social interaction as the outcome variable: (1) *clear* interactions only, to capture precisely defined instances of social interaction, (2) *clear* and *unclear* interactions combined, to include more ambiguous social interaction instances to the set, and (3) all interaction types combined (i.e., including *distant* interactions that took place in some observation sessions). To avoid bias due to class imbalance, I downsampled the number of zero-interaction rows in the training set by stratified undersampling to match the number of rows within positive social interaction class. I did this to prevent the algorithm from learning to classify based on the overrepresentation of instances without social interactions. I computed classification metrics (i.e., accuracy, precision, recall, and F1-score) using the *caret* package (version 7.0.1; Kuhn, 2008) in R (Appendix C, Figure C1). Accuracy refers to the proportion of correctly classified cases, based on a threshold of 0.5 for converting predicted probabilities into class labels. Precision refers to the proportion of correctly predicted positive cases, whereas recall (also known as sensitivity) refers to the proportion of positive cases correctly identified by model. The F1-score is the harmonic mean of precision and recall, thereby balancing false positives and false negatives. In addition to the accuracy metric, I calculated a threshold-independent metric by computing the area under the *receiver operating characteristic curve* (AUC), which reflects how often the model assigns a larger predicted value to an actual interaction than to a non-interaction when comparing all possible pairs. AUC values were computed using the *pROC* package (version 1.18.5; Robin et al., 2011). In order to evaluate whether each predictor, or their combination, had a consistent influence on the classification of social interactions, I collected predictor estimates across test sets for each model. Since both distance and angle were *z*-standardized, their estimates reflected the effect of a one standard deviation change in the predictor. Since the constellation type was not a continuous predictor, I applied Gelman's standardization (i.e., converting 1-unit change of the binary predictor to a 2 SD change in continuous predictor) to the constellation predictor, so that its scale matched that of continuous predictors that have been *z*-standardized (Gelman et al., 2020). Then, I conducted one-sample *t*-tests for the predictor estimates.

I plotted illustrations in MS Excel and in R (version 4.3.0; R Core Team, 2023) using the *ggplot2* package (version 3.5.1; Wickham, 2016) and organized to panels in MS PowerPoint. I further used grouping and counting functions of *tidyr* package (version 1.3.1; Wickham et al., 2024) and *dplyr* package (version 1.1.4; Wickham et al., 2023), and string manipulation functions of the *stringr* packages (version 1.5.0; Wickham, 2023).

# Results

## Descriptive statistics of *F*- and *NF*-constellations

The average distance between dyads was *M* = 0.93 m (*SD* = 0.61, *Mdn* = 0.77). The average distance between children within the *F*-constellation, which only contained befriended dyads, was *M* = 0.94 m (*SD* = 0.63, *Mdn* = 0.77). The average dyadic distance in *NF*-constellation was *M* = 0.77 m (*SD* = 0.57, *Mdn* = 0.57) for befriended dyads, and *M* = 0.99 m (*SD* = 0.60, *Mdn* = 0.85) for non-befriended dyads. The distribution of distances for both constellation types and dyadic relations are displayed in Figure 7.

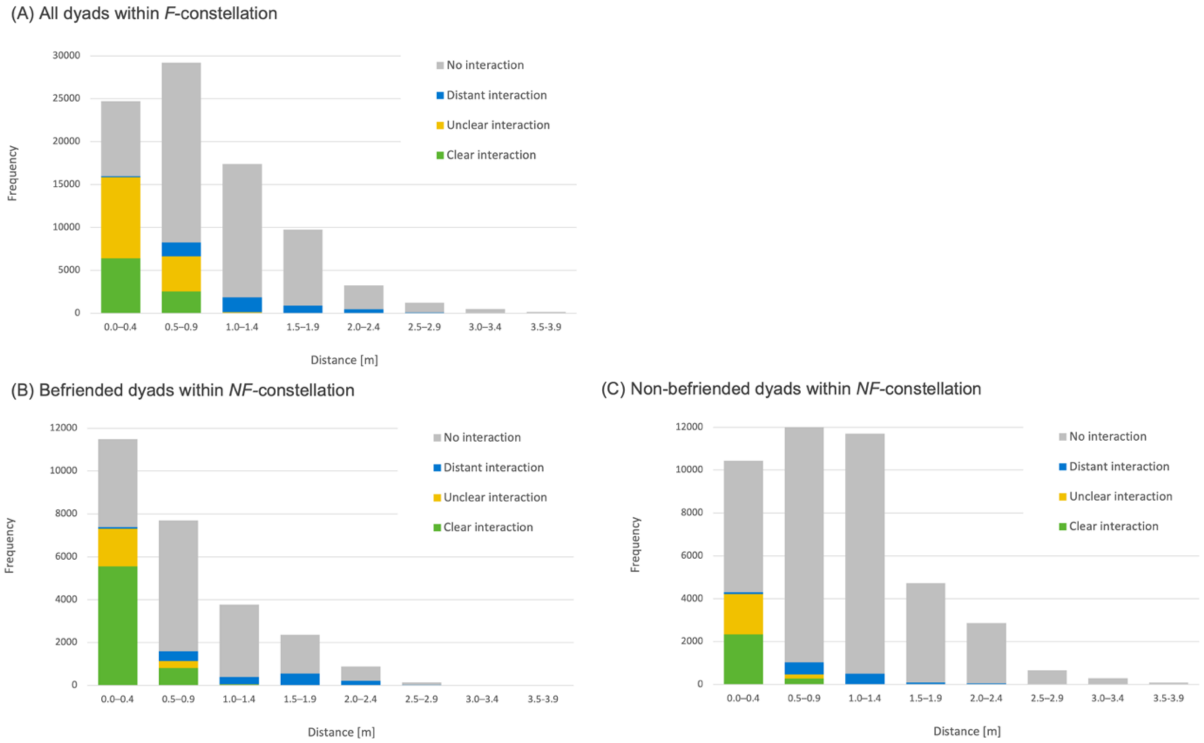



**Figure 7. Histograms of distances across all observations while differentiating between interaction types.** Stacked coloured bars are indicating the absolute frequency of distances (in meters) for each interaction type: dyadic distances at which no social interaction was coded (grey), when *distant* interactions were coded (blue), when *unclear* interactions were coded (yellow), and, finally, distances at which the social interaction was coded as *clear* (green).

The average angle between dyads was *M* = 119.92 degrees (*SD* = 67.27, *Mdn* = 114.83). The average degree for dyads assigned to F-constellation, that contained only befriended dyads, was *M* = 119.51 degrees (*SD* = 68.17, *Mdn* = 113.45). The average degree within NF-constellation was *M* = 115.50 degrees (*SD* = 66.93, *Mdn* = 109.91) for befriended dyads, and *M* = 122.80 degrees (*SD* = 65.73, *Mdn* = 119.63) for non-befriended dyads.

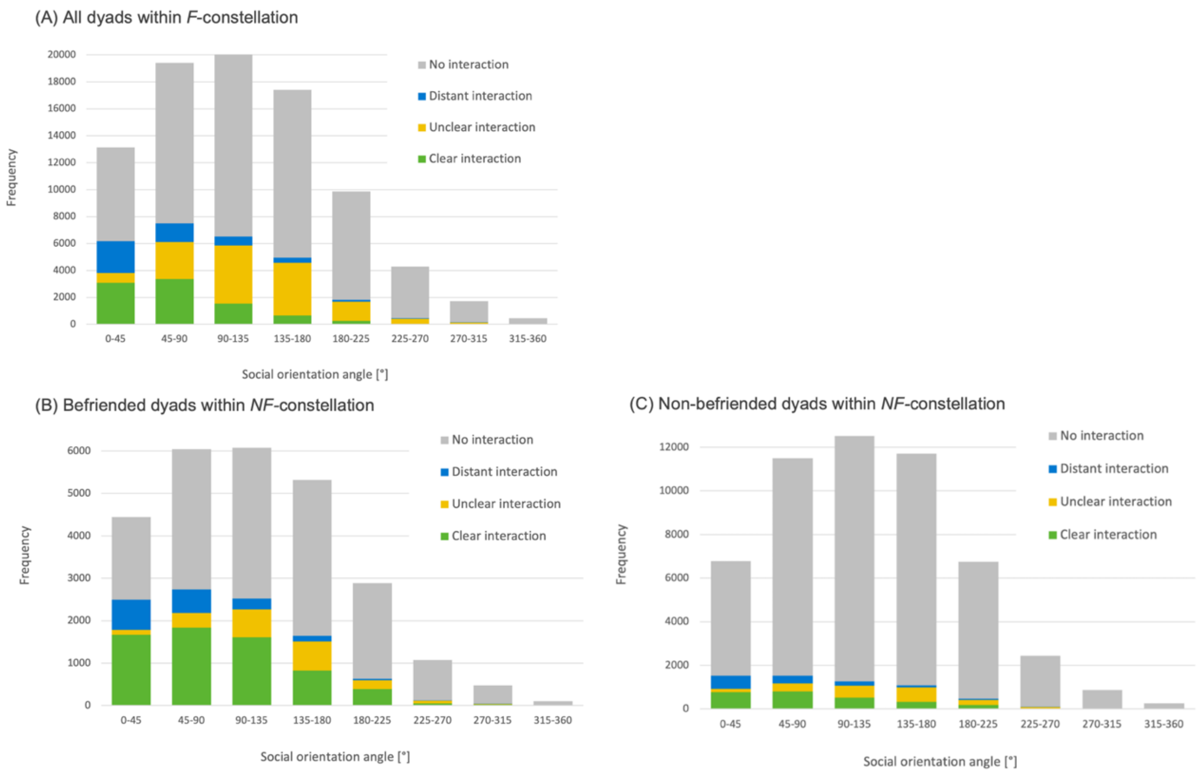



**Figure 8. Histograms of angles across all observations while differentiating between interaction types.** Stacked coloured bars are indicating the absolute frequency of dyadic angles (in degrees) for each interaction type: dyadic distances at which no social interaction was coded (grey), when *distant* interactions were coded (blue), when *unclear* interactions were coded (yellow), and, finally, distances at which the social interaction was coded as *clear* (green).

### The effect of friendship on dyadic distance within the *NF* constellations

I a fitted linear mixed-effects model with random intercepts to investigate the effect of friendship (befriended vs. non-befriended dyads) on dyadic distance, with video observation session included as a random intercept. The model revealed a significant effect of friendship, $b$ = –214.68, SE = 54.15, $t(21)$ = –3.97, $p$ = .001. On average, befriended dyads ($M$ = 0.771 m, $SD$ = 0.309) were approximately 0.215 m closer than non-befriended dyads ($M$ = 0.986 m, $SD$ = 0.309; Figure 9).

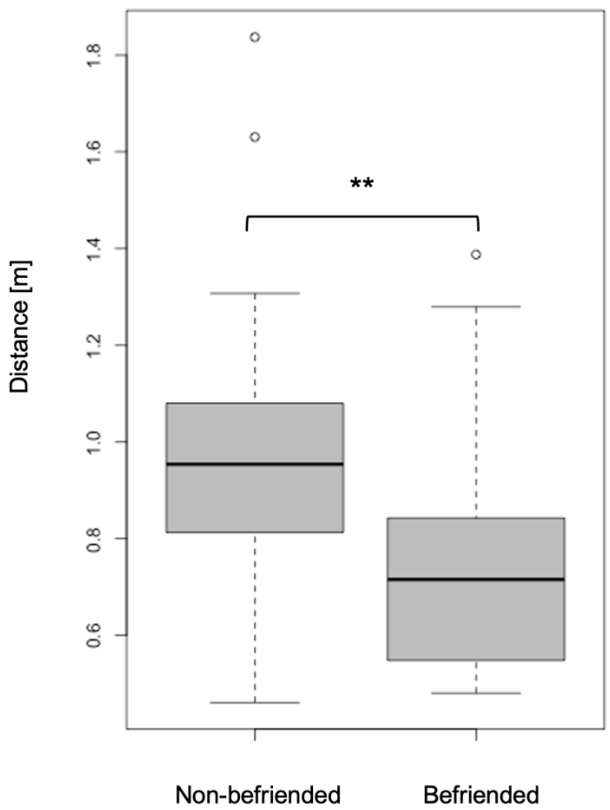



**Figure 9. Boxplot of distances of befriended and non-befriended dyads.** Befriended dyads were significantly closer than non-befriended dyads, p = .001. The horizontal line in each boxplot indicates the median; the boxes represent the interquartile range; whiskers extend to 1.5 × IQR; circles indicate outliers.

Since there were eleven different video observations, I computed Cook's distance to identify the most influential video observations above the threshold of $4/n \approx 0.36$ (Figure 10). The sensitivity analysis after excluding the observation DYAD06NF (D = 0.793) yielded similar results with significant negative effect ($b = -201.53$, SE = 57.41, $t(19) = -3.51$, $p = .002$), suggesting a robust result which was not driven by influence from specific observation.

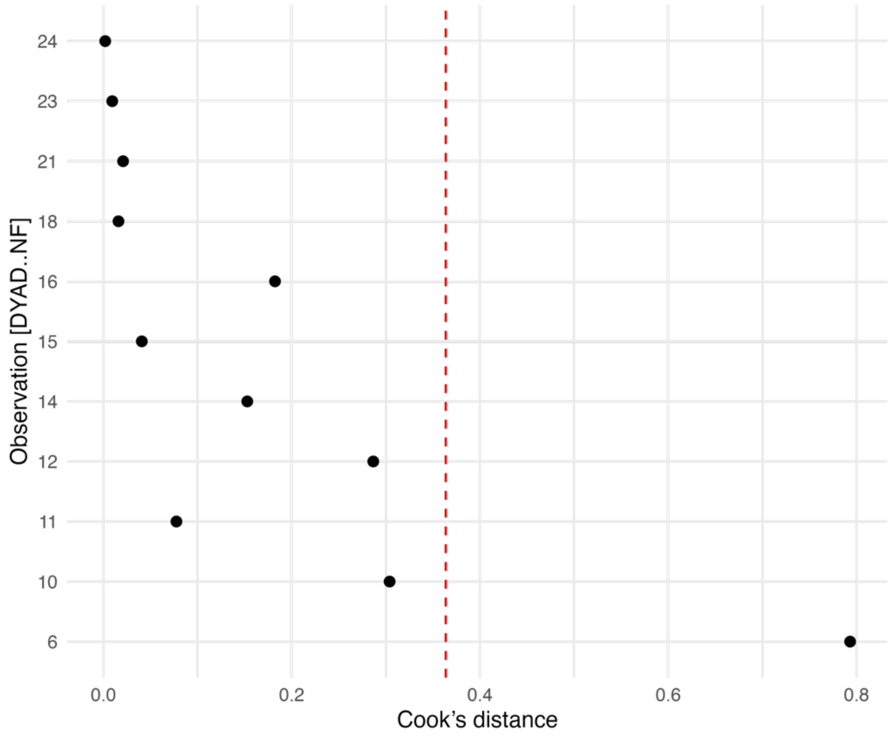



**Figure 10. Cook's distance values across all observations for final model with distance as outcome variable.** Cook's distance of observation DYAD06NF was above the threshold (red dashed line).

**The effect of friendship on dyadic angle within the *NF*-constellations**

To investigate the effect of friendship (befriended vs. non-befriended dyads) on dyadic social orientation angles, I a fitted linear mixed-effects model with random intercepts and included video observation session as random intercept. The model did not reveal a statistically significant effect of friendship on orientation angle, $b$ = –7.30, SE = 4.27, $t(21)$ = –1.71, $p$ = .102. On average, befriended dyads showed by 7.3 degrees more pronounced facing orientation ($M$ = 116.0°, SD = 17.8) than non-befriended dyads ($M$ = 123.0°, SD = 15.0; Figure 11), but the difference was not statistically meaningful.

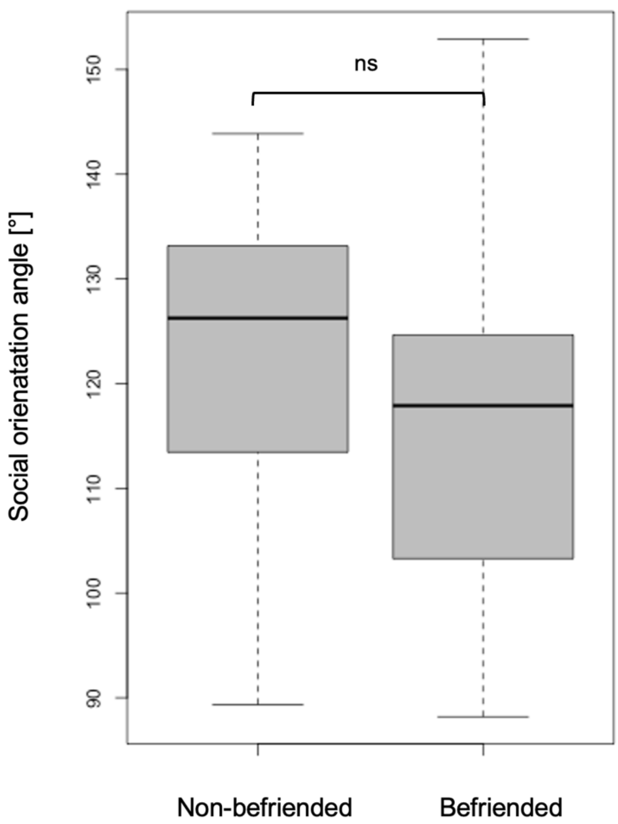



**Figure 11. Boxplot of social orientation angles of befriended and non-befriended dyads.** There was no statistically significant difference between befriended and non-befriended dyads. The horizontal line in each boxplot indicates the median; the boxes represent the interquartile range; whiskers extend to 1.5 × IQR; circles indicate outliers.

I computed Cook's distance to identify the most influential video observations above the threshold of $4/n \approx 0.36$ (Figure 12). Cook's distances of all observations were below the threshold.

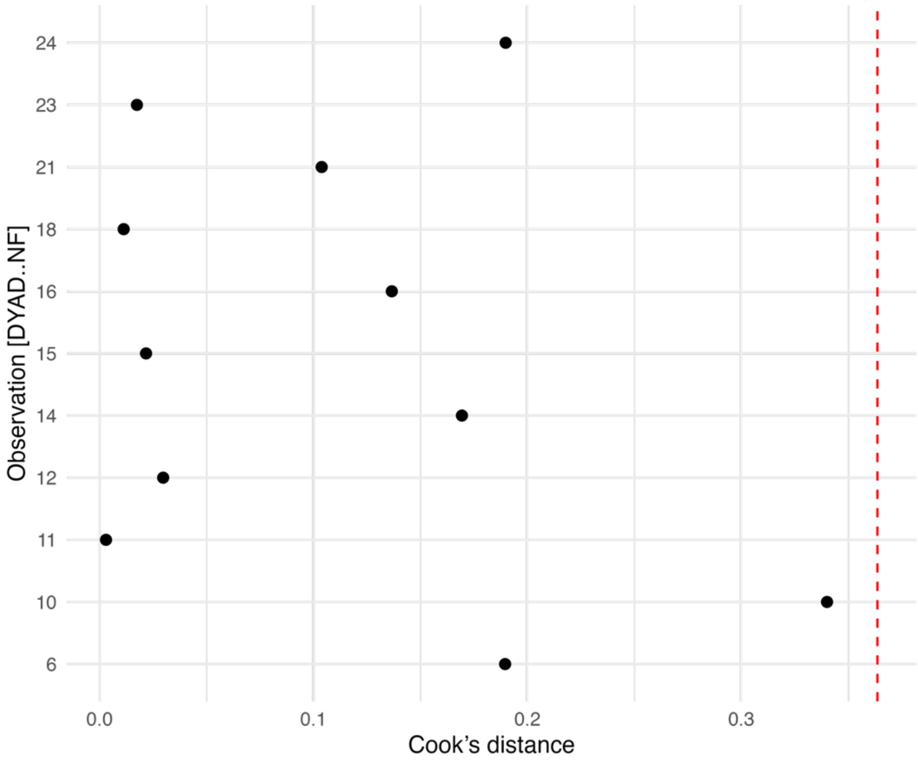



**Figure 12. Cook's distance values across all observations for final model with angle as outcome variable.** None of the Cook's distances exceeded the threshold (red dashed line).

**The effect of friendship on duration of hand-coded social interaction within *NF*-constellations**

To investigate the effect of friendship (befriended vs. non-befriended dyads) on duration of *clear* social interactions, I a fitted linear mixed model with log-transformed *clear* social interactions as outcome and included video observation session as random intercept. The model revealed a significant effect of friendship on duration of *clear* social interactions, $b$ = –2.54, SE = 0.49, $t(21)$ = –5.18, $p$ < .001. On average, befriended dyads ($M$ = 145.0 s, SD = 114.0) had by 115.5 seconds longer *clear* social interactions than non-befriended dyads ($M$ = 29.5 s, SD = 51.8; Figure 13).

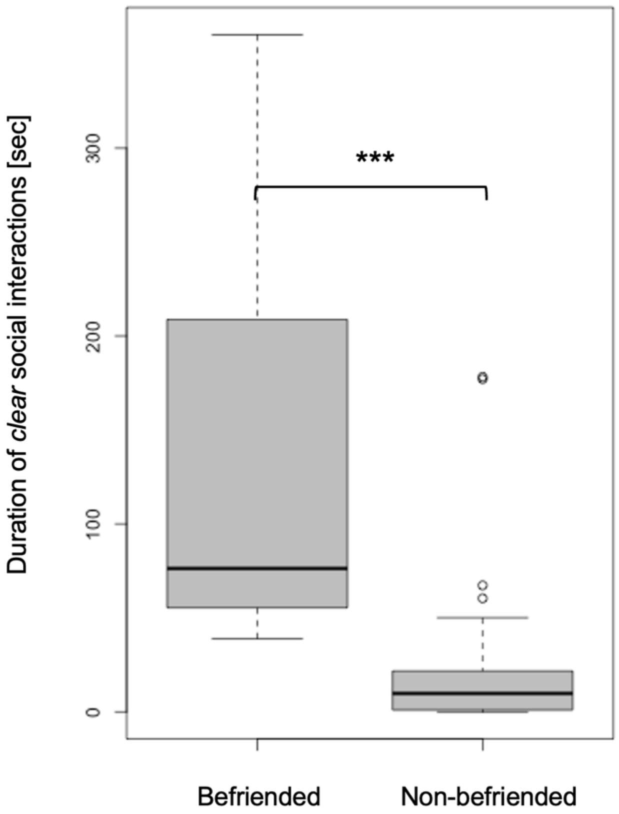



**Figure 13. Boxplot of durations of clear social interactions of befriended and non-befriended dyads.** There was a statistically significant difference between befriended and non-befriended dyads ($p$ < .001). The horizontal line in each boxplot indicates the median; the boxes represent the interquartile range; whiskers extend to 1.5 × IQR; circles indicate outliers.

I computed Cook's distance to identify the most influential video observations above the threshold of $4/n \approx 0.36$ (Figure 14). In the sensitivity analysis after excluding the observation DYAD16NF (D = 0.587), the model revealed a singular fit, suggesting the random intercept added no value (ICC = 0), likely because the variability between observations was low. However, the fixed effect of friendship remained significant and negative ($b$ = –2.74, SE = 0.50, $t$(28) = –5.50, $p$ < .001), indicating robustness of findings.

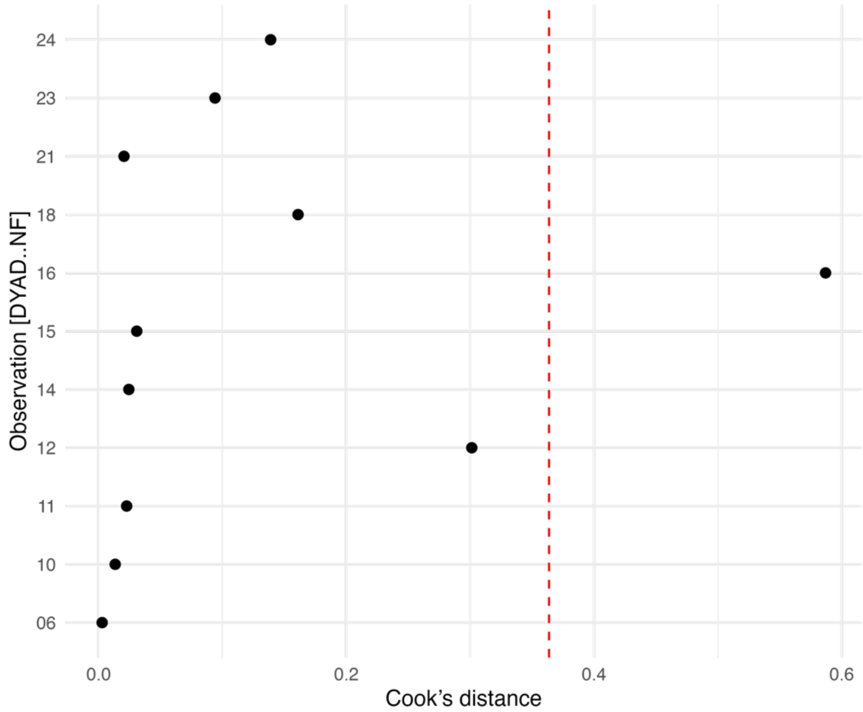



**Figure 14. Cook’s distance values across all observations for final model with distance as outcome variable.** Cook’s distance of observation DYAD16NF was above the threshold (red dashed line).

Since the assumptions of linear mixed models were not fully met, I conducted nonparametric Wilcoxon signed-rank test. Based on median comparisons of 11 paired observations, durations of *clear* interactions were 68.05 seconds longer for befriended dyads (*Mdn* = 76.3 s) than for non-befriended dyads (*Mdn* = 8.25 s), and the result was significant (V = 1, $p$ = .002).

To investigate the effect of friendship (befriended vs. non-befriended dyads) on duration of both *clear* and *unclear* social interactions, I resorted to the nonparametric Wilcoxon signed-rank test due to persisting violations of the homoscedasticity assumption, even after the log-transformation of the outcome. The analysis revealed significant effect of friendship on duration of *clear* and *unclear* interactions (V = 1, $p$ = .002). Befriended dyads engaged significantly longer in this type of interaction (*Mdn* = 159.1 s) compared to non-befriended dyads (*Mdn* = 102.7 s), with a median difference of 102.7 second (Figure 15).

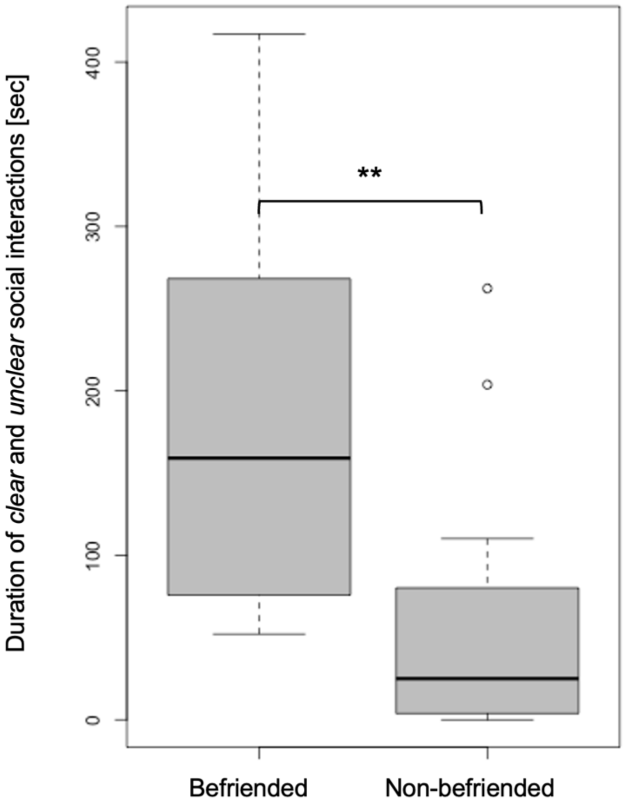



**Figure 15. Boxplot of durations of clear and unclear social interactions of befriended and non-befriended dyads.** There was a statistically significant difference between befriended and non-befriended dyads (*p* = .002). The horizontal line in each boxplot indicates the median; the boxes represent the interquartile range; whiskers extend to 1.5 × IQR; circles indicate outliers.

Finally, I investigated the effect of friendship on all types of coded social interactions combined using nonparametric Wilcoxon signed-rank test due to violations of the homoscedasticity assumption, even after the log-transformation of the outcome. The analysis revealed significant effect of friendship on duration of *clear* and *unclear* interactions (V = 4, *p* = .007). Befriended dyads engaged significantly longer (*Mdn* = 236.7 s, *M* = 230.0, SD = 133.0) in this type of interaction compared to non-befriended dyads (*Mdn* = 33.3 s, *M* = 62.2, *SD* = 88.0), with a median difference of 153.3 seconds (Figure 16).

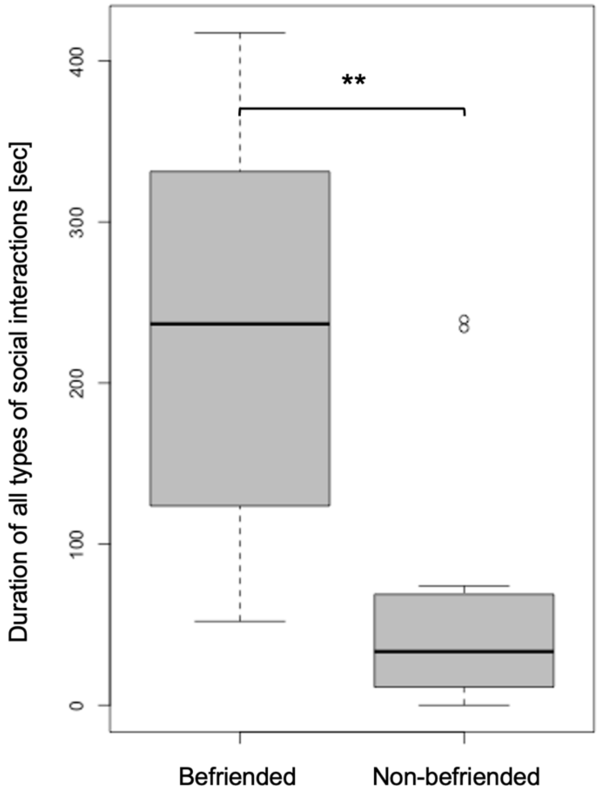



**Figure 16. Boxplot of durations of all types of social interactions of befriended and non-befriended dyads.** There was a statistically significant difference between befriended and non-befriended dyads ($p$ = .007). The horizontal line in each boxplot indicates the median; the boxes represent the interquartile range; whiskers extend to 1.5 × IQR; circles indicate outliers.

**Predicting social interactions using machine learning on data from *F*- and *NF*-constellations**

To investigate the possibility of predicting the instances of social interaction from simple measures of dyadic distance and social orientation angle, I evaluated mixed logistic regression models using LOOCV approach across 23 observations. Table 1 shows summary of performance for all tested models across all outcomes. Overall, performance across models with all outcomes (measured by both accuracy and AUC score) was highest in models with both predictors, and almost equally high in models with distance as single predictor. The overall accuracy and the AUC score of the models relying solely on angle was lowest.

After adding *unclear* interactions to model, the performance of the model with distance as single predictor improved slightly. However, the accuracies and the AUC scores of models including both predictors combined did not improve but dropped slightly. The drop was even more pronounced in the model relying solely on angle. Interestingly, the AUC scores, seemed to remain unchanged independent of whether interaction term was included or not. After adding *distant* interactions to models, the performance of models relying on distance dropped. The performance of the models relying solely on angle, on the other hand, showed minor improvements after including *distant* interactions.

**Table 1. Averaged performance metrics for models across all outcomes**

| | *Clear* | | *Clear-unclear* | | *Clear-unclear-distant* | |
|---|---|---|---|---|---|---|
| Model | Accuracy [%] | AUC | Accuracy [%] | AUC | Accuracy [%] | AUC |
| Distance | 81.26 | 0.910 | 82.79 | 0.908 | 76.80 | 0.839 |
| Angle | 64.30 | 0.703 | 56.98 | 0.616 | 58.12 | 0.639 |
| Distance and angle (additive) | 83.81 | 0.928 | 83.24 | 0.909 | 77.63 | 0.815 |
| Distance and angle (interaction) | 83.70 | 0.928 | 83.30 | 0.909 | 78.23 | 0.839 |

All models were fitted using mixed logistic regression with random intercepts and random slopes for observation as a random effect. Values reflect average test performance from leave-one-out cross-validation across 23 observations. Accuracy is reported as a percentage. AUC = Area Under the receiver operating characteristic Curve. *Clear* represent the average performance across all test sets with *clear* interaction as outcome variable. *Clear-unclear* represent the average performance across all test sets with *clear* and *unclear* interactions combined as outcome variable. All interactions represent the average performance across all test sets with all interaction types combined as outcome variable.

Table 2 shows summary of estimates for all predictors for all tested models across all outcomes. All predictors, including distance, angle, their interaction term, and constellation, showed statistically significant effects ($p < .001$). Distance was the strongest predictor in all models.

**Table 2. Averaged fixed-effect estimates for models across all outcomes**

| Model | Predictor | *Clear* | | *Clear-unclear* | | *Clear-unclear-distant* | |
|---|---|---|---|---|---|---|---|
| | | Est ± *SD* | *t* | Est ± *SD* | *t* | Est ± *SD* | *t* |
| Distance | distance | -5.44 ± 0.18 | -147.47 | -5.05 ± 0.14 | -168.15 | -3.24 ± 0.14 | -97.89 |
| | constellation | 0.55 ± 0.16 | 16.79 | 1.89 ± 0.16 | 57.62 | 1.63 ± 0.14 | 45.50 |
| Angle | angle | -0.88 ± 0.02 | -198.55 | -0.39 ± 0.01 | -134.65 | -0.53 ± 0.14 | -156.49 |
| | constellation | -0.28 ± 0.30 | -4.57 | 1.21 ± 0.27 | 21.26 | 1.24 ± 0.14 | 28.63 |
| Distance and angle (additive) | distance | -5.47 ± 0.18 | -148.53 | -4.98 ± 0.14 | -174.11 | -3.25 ± 0.14 | -98.82 |
| | angle | -1.02 ± 0.03 | -189.31 | -0.35 ± 0.03 | -64.37 | -0.59 ± 0.14 | -128.65 |
| | constellation | 0.30 ± 0.17 | 8.77 | 1.93 ± 0.16 | 57.58 | 1.70 ± 0.14 | 49.28 |
| Distance and angle (interaction) | distance | -5.58 ± 0.18 | -150.37 | -4.97 ± 0.14 | -173.79 | -3.48 ± 0.14 | -109.27 |
| | angle | -1.25 ± 0.05 | -129.18 | -0.29 ± 0.04 | -37.29 | -0.85 ± 0.14 | -139.06 |
| | interaction | -0.29 ± 0.06 | -21.76 | 0.10 ± 0.04 | 11.93 | -0.56 ± 0.14 | -70.79 |
| | constellation | 0.33 ± 0.16 | 9.72 | 1.92± 0.16 | 57.43 | 1.73 ± 0.14 | 51.57 |

All models were fitted using mixed logistic regression with random intercepts and slopes for observation as a random effect. *Clear* represent the average fixed-effect coefficients across all test sets with *clear* interaction as outcome variable. *Clear-unclear* represent the average fixed-effect coefficients across all test sets with *clear* and *unclear* interactions combined as outcome variable. *Clear-unclear-distant* represent the average fixed-effect coefficients across all test sets with all interaction types combined as outcome variable. All estimates were significant ($p < .001$).

For detailed information about performance and averaged fixed-estimates, see Appendix C (Table C1-C6). To illustrate classification report on an example, Figure C1 in Appendix C shows the confusion matrix of the model with distance, angle and constellation as predictors, applied to dataset with DYAD06NF as test set.

## Discussion

In this work, I examined whether simple spatial affiliation cues could be used to identify pre-school children's social interactions in kindergarten settings. Inferential analyses showed that befriended dyads maintained significantly shorter interpersonal distances than non-befriended dyads (by 0.215 m) and engaged in significantly longer interactions across all interaction outcome types. Befriended dyads also exhibited smaller average social orientation angles (by 7.3°), although this effect was not statistically significant. Machine learning models using these cues showed a consistent predictive pattern across all definitions of social interactions. Models including both distance and angle achieved the highest accuracy (~ 83 %) and AUC scores (~ 0.9), followed closely by distance-only models, while angle-only models performed worst. Across all models, predictor estimates were statistically significant, with distance showing the largest negative effect on interaction probability (shorter dyadic distances increasing likelihood of interaction) and social orientation angle showing a smaller negative effect (smaller dyadic angles increasing likelihood of interaction). Taken together, these results indicated strong convergence between inferential and predictive approaches, both identifying proximity as a robust social interaction cue, with social orientation contributing additional explanatory value.

### Effect of friendship on spatial behavior

Within the *NF*-constellation, befriended dyads maintained shorter distances ($M$ = 0.771 m, $Mdn$ = 0.570 m) compared to non-befriended dyads ($M$ = 0.986 m, $Mdn$ = 0.851 m) and the effect was statistically significant ($p$ = .001). Hall (1966) defined the proxemic zone of *personal distance* at 45 to 120 cm for adults, corresponding to the area of sharp vision allowing for discussion of mutual interests. At the same time, Hall (1966) also defined this zone as a distance at which two individuals are able to touch each other when they both stretch out their arms. Since pre-schoolers' arm-span might be approximately 40 to 45 cm (Edmond et al., 2020), one could argue that *personal distance* for children at this age does not exceed 80 to 90 cm. According to this definition, befriended dyads spent most time in *personal distance* proxemic zone defined by two arms' length, while non-befriended dyads spent most time at *social distance* proxemic zone defined by the exceedance of two arms' length. However, Hall's framework was not developed for children and even though no touch was possible for non-befriended dyads at this distance (unless special effort), they remained in the area of sharp vision allowing for discussion of mutual interests, corresponding to the *personal distance* proxemic zone defined by the absolute dyadic distance of 45 to 120 cm. This finding also suggested that personal space at which other peers elicit discomfort and withdrawal (Hayduk, 1987) was slightly bigger within non-befriended dyads than among mutually befriended peers. This was consistent with the finding of Paulus (2008) that pre-school children preferred smaller distances among friends and of Lahnakoski et al. (2020) that adults reported higher enjoyment while interacting in lower proximity.

In the *NF*-constellation, befriended dyads also maintained smaller social orientation angles (*M* = 115.50°, *Mdn* = 109.91°) compared to non-befriended dyads (*M* = 122.80°, *Mdn* = 119.63°), however, the difference was not statistically significant ($p = .102$). Both mean observed social orientation angles (divided by two) would fall into the *peripheral vision* retinal combinations described by Hall (1963), which allow to perceive motion laterally. Relatively large average angles can be explained in the terms of joint attention (Siposova & Carpenter, 2019; Li et al., 2023), when children only briefly entertain eye contact to turn back to mutual toys and maintain *mutual* or *shared attention* during joint play. Since there was no effect of friendship on dyadic angles, the befriended dyads either did not engage in significantly different orientation patterns than non-befriended dyads, contrary to Lahnakoski et al. (2020) who found that orientation towards each other was linked to higher self-reported effort in cooperative tasks in adults.

**Effect of friendship on duration of social interactions**

In the *NF*-constellation, befriended dyads had significantly longer hand-coded social interactions ($p < .01$), regardless of how social interaction was defined (*Mdn* = 76.3 vs. 8.25 sec for *clear* interactions; *Mdn* = 159.1 vs. 102.7 sec for *clear* and *unclear* interactions; and *Mdn* = 236.7 vs. 33.3 sec for all types of interactions combined). This finding aligns with the previous research to interaction length among affiliated children (e.g., Fabes et al, 2003; Howes, 1983; Hartup et al., 1988; Newcomb & Bagwell, 1995) and further supports the use of social interaction instances as an outcome variable to be predicted by distance and angle, which were, in turn, associated with affiliation between children. Interestingly, including *distant* interactions amplified the difference in medians, suggesting that befriended dyads coordinated and engaged with each other across the room, whereas non-befriended dyads tended to stop interacting once they moved apart. The existence of social interactions in *social distance* (up to 360 m) was already highlighted by Hall (1966). However, the pronounced effect in this outcome category could also reflect an artifact of behavioral hand-coding of *clear* and *unclear* interactions, since *distant* interactions (e.g., exchanging objects at a distance) were less ambiguous to categorize and therefore more conservatively annotated than interactions in close proximity.

**Predicting social interactions with machine learning**

I applied machine learning models to spatial measures of dyadic distance and social orientation angle to evaluate their predictive value for identifying social interactions in a naturalistic pre-school setting. All predictors showed statistically significant effects ($p < .001$), suggesting that they are meaningful behavioral features in this context. Across all outcome types, the best-performing models in terms of both accuracy (~ 83 %) and AUC scores (~ 0.9) were those that included both distance, angle and their interaction as predictors. In these models, the negative estimates of distance were also slightly amplified (-5.58 vs. -5.44), likely because including angle and the interaction term allowed the model to better capture multi-dimensional social cues (e.g., aligned orientation at close distance). This finding is consistent with Dai et al. (2020), who used distance and orientation measures from proximity sensors to reconstruct social interactions with high accuracy. Similarly, Setti et al. (2015) showed that proximity and orientation features can be used to identify conversational groups, whereas Lahnakoski et al. (2020) found association with perceived quality of social interactions. However, models relying solely on dyadic distance performed almost equally well as models combining both features, suggesting that proximity alone is a highly reliable cue for identifying social interactions in this age group. Distance had the biggest and the most negative effect estimate on predicting interaction class, with shorter distances increasing probability of social interactions. This finding is consistent with Siposova and Carpenter (2019) who argued that proximity is often necessary for sustained shared attention. Gaze information, on the other hand, is rather interwoven with other cues and should be interpreted in a broader context of social interactions (Hessels, 2020).

Overall, the predictive value of spatial features varied depending on how social interaction was defined. When *unclear* interactions were added to *clear* interactions, the distance coefficient dropped in magnitude and the constellation coefficient became more positive. Since *unclear* interactions were mostly occurring in *F*-constellations, group composition gained more explanatory power than proximity. When *distant* interactions were included in the outcome, performance of distance-only model dropped considerably. This reflected the fact that some interactions took place at longer ranges (e.g., throwing or calling), reducing the relevance of close proximity as a signal for social interaction.

This finding is consistent with proxemics (Hall, 1966) which recognizes social interactions not only at *personal* distance but also at *social distance* (up to 360 cm).

The performance of models relying solely on the social orientation angle was lowest of all predictor combinations (accuracy ≈ 60 %; AUC ≈ 0.65). Also, the average estimates of angle as predictor across all models was distinctly weaker than average estimates of distance. Still, smaller orientation angles (i.e., facing each other) increased the probability of interactions, consistent with neuroimaging evidence from Drijvers & Holler (2022) highlighting the role of face-to-face orientation in social interactions. However, not all meaningful social interactions require direct eye contact and face-to-face orientation but can be explained by *mutual* or *shared* attention during mutual play (Siposova & Carpenter, 2019). Hessels (2020) pointed out that gaze information is closely tight to broader interactive context and shall be interpreted accordingly. Children may often disengage and re-engage in gaze alignment during common activities, as already reflected in the behavioral coding of social interaction (aligned orientation was indicative of continuous social interaction despite disengagement and re-engagement within a time interval of 10 seconds).

The predictive power of angle was highest for predicting *clear* interactions, lower for predicting all types of interactions and worst for predicting *clear-unclear* interactions. After adding *unclear* interactions to outcome, children might have been interacting without heads being oriented towards each other and became less consistently aligned when engaging, possibly maintaining *shared attention* (Siposova & Carpenter, 2019). However, some patterns were introduced reflecting situations where children are interacting despite being poorly aligned (e.g., shared activity with indirect gaze, side-by-side play), which the model could still use for classification. When including *distant* interactions to model, the estimates for angle were still higher compared to *clear* and *unclear* interactions, suggesting that face-to-face alignment became relatively more important after including long-distance engagement, such as calling or throwing objects.

The interaction term between distance and angle had the largest negative estimate after including *distant* interactions to outcome, indicating that the predictive power of angle declined as distance increased. This contrasts with Argyle and Dean’s (1965) finding that closer proximity is associated with reduced eye contact, and with Hall’s (1966) observation that during conversations occurring in *social distance* it is more important to maintain eye contact than at closer proximity. However, Hall’s framework was developed for adults and focused on verbal exchange at the distance.

Children in pre-school age likely did not spend much time in discussions which would require eye contact, but they rather engaged in games. The importance of social orientation angle might have been inhibited due to characteristics of distant play. For example, while jointly manipulating same objects, the attention switched between toys and game partner. Also, throwing objects to each other on distance includes longer periods of searching for non-successfully caught object, especially in the pre-school age with motoric skills still evolving (e.g., Bakeman & Adamson, 1984).

The predictor constellation (i.e., friendship structure within the triad) also contributed to model performance. Being in a friendship triad (*F*-constellation) generally increased the likelihood of interaction. The exception was the angle-only model for *clear* social interactions, where being in the *F*-constellation was associated with lower predicted probability of interaction. The association was weak but might indicate that dyads in *NF*-constellations who did interact may have done so through more directly oriented behaviors, whereas dyads in *F*-constellation may have interacted in side-by-side or less overtly aligned ways, possibly switching between more interaction partners at once.

**Theoretical and practical implications**

Recent studies have demonstrated the value of wearable sensing technology for capturing detailed proximity and orientation data in naturalistic social environments, particularly for populations at risk of social exclusion or atypical interaction patterns. For example, Drye et al. (2024) used ultra-wideband radio frequency identification tracking to record location and orientation information from multiple children with and without autism spectrum disorder, enabling objective calculation of social approach velocities and time spent in social contact. Similarly, Eichengreen et al. (2024) examined schoolyard interactions among deaf and hard-of-hearing preadolescents and their hearing peers using lightweight proximity sensor badges, worn on the chest to detect when two students were near and oriented toward each other. While such approaches provide objective measures of social behavior, incorporating machine learning could extend these methods by automated classification of interaction events and the evaluation of the relative predictive value of different spatial cues. Improved engagement monitoring would have implications for early educational interventions.

Additionally, by using simple spatial cues to describe social interactions, Hall's framework (1963, 1966) might be expanded to children. While auditory and visual abilities are comparable to adults, reaching radius, the size of personal space (Hayduk, 1987) and engagement in different behaviors than adults might lead to discrepancies.

Last but not least, basic markers of affiliation could be universal for other vertebrates (e.g., mammals and birds; Fraser & Bugnyar, 2010). Automated tracking of behaviors can speed up research on joint attention (Siposova and Carpenter, 2019) in non-human animals, and thereby contribute to the ongoing debate about perspective, knowledge, belief and desire attribution in non-human animals (Dennett, 1988).

**Limitations**

The present study was limited by a relatively small sample size. Therefore, children were observed as members of multiple groups to obtain sufficient interaction data, causing within-group dependencies. However, since both tracking and interaction detection can be automated, the methodology is easily scalable, so increase in sample size is quite straightforward in future research.

Because the data are time series, consecutive data points are not independent, and therefore the reported p-values should be interpreted with caution. Nevertheless, this issue affects only the coefficient estimates and their *p*-values, which are computed immediately after fitting the model to the training sets, whereas the performance metrics were computed on unseen test sets and remain unaffected. One way to address this limitation is to model temporal autocorrelation by including lagged predictors.

Possible limitation was the inconsistency in manual coding of social interactions across different parts of the video frame. This inconsistency was caused by the combination of peripheral distortion artifact and the definition of *clear* and *unclear* social interactions, which required an estimation of reaching distance, approximated as the distance at which children could touch with fingertips if both extended their arms (Hall, 1966). While the human coder might have been able to compensate for the distortion, the automatic tracking system relied on the custom peripheral distortion correction function applied to Loopy coordinates based on the height of the table. However, it was only a post-hoc approximation of the 3D space at the level of sitting children, not accounting for variation in children’s posture (standing vs. crawling). Improved calibration using a grid placed along the room’s walls, combined with depth-sensitive multi-camera array, would allow for more accurate approximation of the 3D space.

When the front marker was occluded, the automated tracking software Loopy mistook the opening at the back of the colored cap (especially in combination the black hair color) for the front hat marker, ultimately leading to the opposite orientation estimates. Training of the detection algorithm for the described scenario, or identification and subsequent exclusion of the abrupt switches in head orientation could resolve this issue. Moreover, setting up a multi-camera array would not only minimize occlusion, but also help capture missing cues like gaze and affect, or prevent misclassification of parallel play (e.g., sharing a toy basket) as social interaction.

**Future directions**

An interesting avenue would be to investigate differences in spatial measures among befriended dyads depending on their affiliative relationship to the third child in the room. My descriptive statistics indicated differences in distances in focus dyads depending on the constellation type, with children from the focus dyad maintaining closer proximity among themselves when the third child was mutually disliked. Studies to group composition could, for example, contribute to the person–other–object (P-O-X) framework of the triadic balance theory (Heider, 1958; Cartwright & Harary, 1956), which may link physical spacing to attitudes within triads. In the balance theory, a situation where two friends (P and O) both dislike the same third person (X) creates a balanced triad, where two negatives (P>X and O>X) and one positive (P>O) add to a positive product. The balanced state reinforces the positive tie between P and O at shared bond over the dislike of X, which would strengthen the perception of closeness and reduce distance between the mutual friends.

Previous research has suggested that there may be temporo-spatial constructs associated with affiliation, such as velocity of approach and withdrawal (e.g., Banarjee et al., 2023) or joint spatial dynamics (e.g., Xu et al., 2020). Although the topic of interpersonal synchrony has been well researched (Feldman, 2007), coordination of spatial positioning and locomotion below school age remains underexamined to date. Among the few studies in this area, Hoch et al. (2021) reported that toddlers spontaneously match their movement patterns to those of their mothers during free play. Introducing automated collection and classification of joint spatial dynamics' markers (i.e., Hudson et al., 2023) could enhance research in kindergarten settings.

Future research could also combine various tracking methods with machine learning to advance automated detection and classification of pre-schoolers' social interactions. For example, the UWB RFID tracking method require no line-of-sight, with signal penetrating non-metal obstacles (table, toys, other peers), allowing for consistent coverage in cluttered environments, such as kindergarten classrooms. This approach has already been successfully applied to capture detailed social approach velocity and time in social contact in pre-school classrooms (Drye et al., 2024).

Furthermore, a multi-layered approach (see Horn et al., 2024) using video-based markers of toys, hand position and shoulder orientations (e.g., via colored tags or wearable devices) or automatic pose estimation (Cao et al., 2021), alongside audio recorders with automated speech recognition could provide further information to yield insights into the complexity of pre-school children's interactions. When coupled with machine learning models, unobstructed audio and high-resolution temporo-spatial data could be used not only to describe but also to predict social interactions in real time.

**Conclusion**

This work shows that simple spatial cues, especially proximity, can indicate social interactions in pre-schoolers. Applying of machine learning to large behavioral data could become a precise, time-efficient tool for classification of social interactions.

## References


Afshordi, N., & Liberman, Z. (2021). Keeping friends in mind: Development of friendship concepts in early childhood. *Social Development, 30*(2), 331–342. https://doi.org/10.1111/sode.12493

Ahmadi, M. N., Pavey, T. G., & Trost, S. G. (2020). Machine learning models for classifying physical activity in free-living pre-school children. *Sensors, 20*(6), 4364. https://doi.org/10.3390/s20164364

Aiello, J. R., & De Carlo Aiello, T. (1974). The development of personal space: Proxemic behavior of children 6 through 16. *Human ecology*, *2*(3), 177-189. https://www.jstor.org/stable/4602298

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Archie, E. A., Tung, J., Clark, M., Altmann, J., & Alberts, S. C. (2014). Social affiliation matters: Both same-sex and opposite-sex relationships predict survival in wild female baboons. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1793), 20141261. https://doi.org/10.1098/rspb.2014.1261

Argyle, M., & Dean, J. (1965). Eye-Contact, Distance and Affiliation. *Sociometry, 28*(3), 289–304. https://doi.org/10.2307/2786027

Bakeman, R., & Adamson, L. B. (1984). Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child development*, 1278-1289.

Banarjee, C., Tao, Y., Fasano, R. M., Song, C., Vitale, L., Wang, J., Shyu, M.-L., Perry, L. K., & Messinger, D. S. (2023). Objective quantification of homophily in children with and without disabilities in naturalistic contexts. *Scientific Reports 903*(13). https://doi.org/10.1038/s41598-023-27819-6

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Boucherie, P. H., Mariette, M. M., Bret, C., & Dufour, V. (2016). Bonding beyond the pair in a monogamous bird: impact on social structure in adult rooks (Corvus frugilegus). *Behaviour*, *153*(8), 897-925. https://doi.org/10.1163/1568539X-00003372

Bradski, G. (2000). The opencv library. *Dr. Dobb's Journal: Software tools for the professional programmer*, *25*(11), 120-123. (Version 4.10.0) [Python package]. https://github.com/opencv/opencv

Campos, J. J., Anderson, D. I., Barbu-Roth, M. A., Hubbard, E. M., Hertenstein, M. J., & Witherington, D. (2000). Travel Broadens the Mind. *Infancy*, *1*(2), 149–219. https://doi.org/10.1207/S15327078IN0102_1

Cao, Z., Hidalgo, G., Simon, T., Wei, S.-E., & Sheikh, Y. (2021). OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 43*(1), 172–186. https://doi.org/10.1109/TPAMI.2019.2929257

Cartwright, D., & Harary, F. (1956). Structural balance: a generalization of Heider's theory. *Psychological Review, 63*(5), 277–293. https://doi.org/10.1037/h0046049

Connor, R. C., Heithaus, M. R., & Barre, L. M. (1999). Superalliance of bottlenose dolphins. *Nature, 397*(6720), 571–572. https://doi.org/10.1038/17501

Cords, M., & Aureli, F. (2000). Reconciliation and relationship qualities. In F. Aureli & F. B. M. de Waal (Eds.), *Natural conflict resolution* (pp. 177–198). University of California Press. https://doi.org/10.1525/9780520924932-014

Cristani, M., Bazzani, L., Paggetti, G., Fossati, A., Tosato, D., Del Bue, A., Menegaz, G., & Murino, V. (2011). Social interaction discovery by statistical analysis of F-formations. In D. Metaxas, L. Quan, A. Sanfeliu, & L. Van Gool (Eds.), *Proceedings of the British Machine Vision Conference (BMVC 2011)* (pp. 23.1–23.12). BMVA Press. https://doi.org/10.5244/C.25.23

Dai, S., Bouchet, H., Nardy, A., Fleury, E., Chevrot, J. P., & Karsai, M. (2020). Temporal social network reconstruction using wire-less proximity sensors: Model selection and consequences. *EPJ Data Science, 9*(19). https://doi.org/10.1140/epjds/s13688-020-00237-8

Dennett, D. C. (1988). Précis of the intentional stance. *Behavioral and brain sciences, 11*(3), 495-505. https://doi.org/10.1017/S0140525X00058611

Drijvers, L., & Holler, J. (2022). Face-to-face spatial orientation fine-tunes the brain for neurocognitive processing in conversation. *Iscience, 25*(11). https://doi.org/10.1016/j.isci.2022.105413

Drye, M., Banarjee, C., Perry, L., Viggiano, A., Irvin, D., & Messinger, D. (2024). Children's social preference for teachers versus peers in autism inclusion classrooms: An objective perspective. *Autism Research.* 18(1), 179-194. https://doi.org/10.1002/aur.3276

Dunbar, R. I. M. (2012). Bridging the bonding gap: The transition from primates to humans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1597), 1837–1846. https://doi.org/10.1098/rstb.2011.0217

Edmond, T., Laps, A., Case, A. L., O'Hara, N., & Abzug, J. M. (2020). Normal Ranges of Upper Extremity Length, Circumference, and Rate of Growth in the Pediatric Population. *Hand (New York, N.Y.), 15*(5), 713–721. https://doi.org/10.1177/1558944718824706

Eichengreen, A., Tsou, Y.-T., van Klaveren, L.-M., Zaidman-Zait, A., & Koutamanis, A. (2024). Using wearable sensors to explore schoolyard interactions of mainstreamed deaf and hard-of-hearing preadolescents. Social Development. Advance online publication. https://doi.org/10.1111/sode.12755

Erdley, C. A., & Day, H. J. (2017). *Friendship in childhood and adolescence.* In A. L. Weber & C. H. Johnson (Eds.), *The psychology of friendship* (pp. 4-11). Oxford University Press.

Estes, B. W., & Rush, D. (1971). Social schemas: A developmental study. *Journal of Psychology, 78*(1), 119–123.

Fabes, R. A., Martin, C. L., & Hanish, L. D. (2003). Young children's play qualities in same-, other-, and mixed-sex peer groups. *Child Development, 74*(3), 921–932. https://doi.org/10.1111/1467-8624.00576

Feldman, R. (2007). Parent-infant synchrony and the construction of shared timing; physiological precursors, developmental outcomes, and risk conditions. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 48*(3–4), 329–354. https://doi.org/10.1111/j.1469-7610.2006.01701.x

Fischer, C. S. (1968). Social schemas: Response sets or perceptual meanings? *Journal of Personality and Social Psychology, 10*(1), 8–14.

Fox, J., & Weisberg, S. (2019). *Car: Companion to applied regression* (Version 3.1-2) [R package]. https://CRAN.R-project.org/package=car

Fraser, O. N., & Bugnyar, T. (2010). The quality of social relationships in ravens. *Animal Behaviour, 79*(4), 927–933. https://doi.org/10.1016/j.anbehav.2010.01.008

Fraser, O. N., & Bugnyar, T. (2012). Reciprocity of agonistic support in ravens. *Animal Behaviour*, *83*(1), 171-177. https://doi.org/10.1016/j.anbehav.2011.10.023c

Fraser, O. N., Schino, G. & Aureli, F. (2008). Components of relationship quality in chimpanzees. *Ethology, 114*, 834–843. https://doi.org/10.1111/j.1439-0310.2008.01527.x

Gelman, A., Hill, J., &; Vehtari, A. (2020). Regression and Other Stories. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781139161879

Hagenbuchner, M., Cliff, D. P., Trost, S. G., Van Tuc, N., & Peoples, G. E. (2015). Prediction of activity type in pre-school children using machine learning techniques. *Journal of Science and Medicine in Sport, 18*(4), 426–431. https://doi.org/10.1016/j.jsams.2014.06.003

Hall, E. T. (1963). A System for the Notation of Proxemic Behavior. *American Anthropologist, 65*(5), 1003–1026. http://www.jstor.org/stable/668580

Hall, E. T. (1966). The hidden dimension. Garden City, NY: Doubleday.

Hall, J. A. (2012). Friendship standards: The dimensions of ideal expectations. *Journal of Social and Personal Relationships, 29*(7), 884-907. https://doi.org/10.1177/0265407512448274

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., … & Oliphant, T. E. (2020). Array programming with NumPy. *Nature, 585*(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Hartig, F. (2022). *DHARMa: Residual diagnostics for hierarchical (multi-level / mixed) regression models*. (Version 0.4.7) [R package]. https://CRAN.R-project.org/package=DHARMa

Hartup, W. W., Laursen, B., Stewart, M. I., & Eastenson, A. (1988). Conflict and the friendship relations of young children. *Child Development, 59*(6), 1590–1600. https://doi.org/10.2307/1130673

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning*. Springer. https://doi.org/10.1007/978-0-387-84858-7

Hayduk, L. A. (1983). Personal space: Where we now stand. *Psychological Bulletin, 94*(2), 293–335. https://doi.org/10.1037/0033-2909.94.2.293

Hayduk, L. A. (1978). Personal space: An evaluative and orienting overview. *Psychological Bulletin, 85*(1), 117–134. https://doi.org/10.1037/0033-2909.85.1.117

Hediger, H. (1950). *Wild animals in captivity*. London, England: Butterworth & Company.

Hessels, R. S. (2020). How does gaze to faces support face-to-face interaction? A review and perspective. *Psychonomic bulletin & review, 27*(5), 856-881. https://doi.org/10.3758/s13423-020-01715-w

Heider, F. (1958). *The psychology of interpersonal relations.* John Wiley & Sons Inc. https://doi.org/10.1037/10628-000

Hinde, R. A. (1976). Interactions, relationships and social structure. *Man, 11*(1), 1–17. https://doi.org/10.2307/2800384

Hoch, J. E., Ossmy, O., Cole, W. G., Hasan, S., & Adolph, K. E. (2021). “Dancing” Together: Infant–Mother Locomotor Synchrony. *Child Development, 92*(4), 1337–1353. https://doi.org/10.1111/cdev.13513

Horn, L., Karsai, M., & Markova, G. (2024). An automated, data-driven approach to children’s social dynamics in space and time. *Child Development Perspectives*. Advance online publication. https://doi.org/10.1111/cdep.12495

Howes, C. (1983). Patterns of friendship. *Child Development, 54*(4), 1041–1053. https://doi.org/10.2307/1129908

Hudson, D., Wiltshire, T.J. & Atzmueller, M. (2023). multiSyncPy: A Python package for assessing multivariate coordination dynamics. *Behavioral Research, 55*, 932–962. https://doi.org/10.3758/s13428-022-01855-y

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in science & engineering, 9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Johnson-Ulrich, L. (2017). The Social Intelligence Hypothesis. In T. K. Shackelford & V. A. Weekes-Shackelford (Eds.), *Encyclopedia of Evolutionary Psychological Science* (pp. 1–7). Springer. https://doi.org/10.1007/978-3-319-16999-6_3100-1

Katz, N. (1937). *Animals and men.* New York, NY: Longmans, Green.

Kendon, A. (1977). *Studies in the behavior of face-to-face interaction*. Lisse, The Netherlands: Peter de Ridder Press.

Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of statistical software, 28*(5), 1–26. https://doi.org/10.18637/jss.v028.i05

Kutsukake, N. (2009). Complexity, dynamics and diversity of sociality in group-living mammals. Ecological Research, 24(4), 521–531. https://doi.org/10.1007/s11284-008-0563-4

Lahnakoski, J. M., Forbes, P. A., McCall, C., & Schilbach, L. (2020). Unobtrusive tracking of interpersonal orienting and distance predicts the subjective quality of social interactions. *Royal Society Open Science*, *7*(8), 191815. http://dx.doi.org/10.1098/rsos.191815

Li, P., Lu, H., Poppe, R. W., & Salah, A. A. (2023, October). Automated detection of joint attention and mutual gaze in free play parent-child interactions. In *Companion Publication of the 25th International Conference on Multimodal Interaction* (pp. 374-382). ACM. https://doi.org/10.1145/3610661.3616234

Loth, A. (2017). Individual recognition in bottlenose dolphins (Tursiops truncatus): From acoustic distinctiveness in signature whistles to visual self-recognition. (Doctoral dissertation). University of St Andrews. https://doi.org/10.17630/10023-12097

Luef, E. M., Ter Maat, A., & Pika, S. (2017). Vocal similarity in long-distance and short-distance vocalizations in raven pairs (Corvus corax) in captivity. *Behavioural Processes*, *142*, 1-7. https://doi.org/10.1016/j.beproc.2017.05.013

McBride, G. (1964). *A general theory of social organization and behavior*. St. Lucia, Australia: University of Queensland Press.

McBride, G., King, M. G., & James, J. W. (1965). Social proximity effects on galvanic skin response in adult humans. *Zeitschrift für Psychologie, 171*(3), 153–157.

McFarland, R., & Majolo, B. (2011). Grooming Coercion and the Post-Conflict Trading of Social Services in Wild Barbary Macaques. *PLoS ONE, 6*(10), e26893. https://doi.org/10.1371/journal.pone.0026893

McKinney, W. (2010). Data structures for statistical computing in Python. In S. van der Walt & J. Millman (Eds.), *Proceedings of the 9th Python in Science Conference* (pp. 56–61). https://doi.org/10.25080/Majora-92bf1922-00a

Mehrabian, A. (1968). Relationship of attitudes to seated posture, orientation, and distance. *Journal of Personality and Social Psychology, 10*(1), 26–30. https://doi.org/10.1037/h0026384

Messinger, D. S., Prince, E. B., Zheng, M., Martin, K., Mitsven, S. G., Huang, S., Stölzel, T., Johnson, N., Rudolph, U., Perry, L. K., Laursen, B., & Song, C. (2019). Continuous measurement of dynamic classroom social interactions. *International Journal of Behavioral Development*, *43*(3), 263–270. https://doi.org/10.1177/0165025418820708

Mukhopadhyay, S. C. (2014). Wearable sensors for human activity monitoring: A review. *IEEE Sensors journal, 15*(3), 1321–1330. https://doi.org/10.1109/JSEN.2014.2370945

Newcomb, A. F., & Bagwell, C. L. (1995). Children's friendship relations: A meta-analytic review. *Psychological Bulletin, 117*(2), 306–347. https://doi.org/10.1037/0033-2909.117.2.306

Osmond, H. (1957). *Function as the basis of psychiatric ward design*. Psychiatric Services, 8(4), 23–27. https://doi.org/10.1176/ps.8.4.23

Paulus, M. (2018). Pre-school children's and adults' expectations about interpersonal space. *Frontiers in Psychology, 9*, 2479. https://doi.org/10.3389/fpsyg.2018.02479

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825–2830. https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf

Python Software Foundation. (2023). *Python (Version 3.13.0) [Computer software]*. https://www.python.org

R Core Team. (2023). *R: A language and environment for statistical* computing (Version 4.3.0) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., & Müller, M. (2011). pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics, 12*, 77. https://doi.org/10.1186/1471-2105-12-77

Samuni, L., Preis, A., Mielke, A., Deschner, T., Wittig, R. M., & Crockford, C. (2018). Social bonds facilitate cooperative resource sharing in wild chimpanzees. *Proceedings of the Royal Society B*, *285*(1888), 20181643. https://doi.org/10.1098/rspb.2018.1643

Scheiber, I. B., Kotrschal, K., & Weiß, B. M. (2009). Serial agonistic attacks by greylag goose families, *Anser anser,* against the same opponent. *Animal behaviour*, *77*(5), 1211-1216. https://doi.org/10.1016/j.anbehav.2009.01.026

Scheid, C., Range, F., & Bugnyar, T. (2007). When, what, and whom to watch? Quantifying attention in ravens (*Corvus corax*) and jackdaws (*Corvus monedula*). *Journal of Comparative Psychology, 121*(4), 380–386. https://doi.org/10.1037/0735-7036.121.4.380

Schielzeth, H., Dingemanse, N. J., Nakagawa, S., Westneat, D. F., Allegue, H., Teplitsky, C., Réale, D., Dochtermann, N. A., Garamszegi, L. Z., & Araya-Ajoy, Y. G. (2020). Robustness of linear mixed-effects models to violations of distributional assumptions. *Methods in Ecology and Evolution*, *11*(9), 1141–1152. https://doi.org/10.1111/2041-210X.13434

Schino, G. (2007). Grooming and agonistic support: A meta-analysis of primate reciprocal altruism. *Behavioral Ecology, 18*(1), 115–120. https://doi.org/10.1093/beheco/arl045

Schwab, C., Bugnyar, T., Schloegl, C., & Kotrschal, K. (2008). Enhanced social learning between siblings in common ravens (*Corvus corax*). *Animal Behaviour, 75*(2), 501–508. https://doi.org/10.1016/j.anbehav.2007.06.006

Selman, R. L. (1981). The development of interpersonal competence: The role of understanding in conduct. *Developmental Review, 1*(4), 401–422. https://doi.org/10.1016/0273-2297(81)90034-4

Setti, F., Russell, C., Bassetti, C., & Cristani, M. (2015). F-formation detection: Individuating free-standing conversational groups in images. *PLOS ONE, 10*(5), e0123783. https://doi.org/10.1371/journal.pone.0123783

Silk, J. B. (2007). Social components of fitness in primate groups. *Science*, 317(5843), 1347–1351. https://doi.org/10.1126/science.1140734

Siposova, B., & Carpenter, M. (2019). A new look at joint attention and common knowledge. *Cognition, 189*, 260–274. https://doi.org/10.1016/j.cognition.2019.03.019

Spector, R. H. (1990). Visual fields. In H. K. Walker, W. D. Hall, & J. W. Hurst (Eds.), *Clinical methods: The history, physical, and laboratory examinations* (3rd ed., chap. 116). Boston, MA: Butterworths.

Sussman, R. W., Garber, P. A., & Cheverud, J. M. (2005). Importance of cooperation and affiliation in the evolution of primate sociality. *American Journal of Physical Anthropology, 128(1), 84–97.* https://doi.org/10.1002/ajpa.20196

Szipl, G., Boeckle, M., Wascher, C. A., Spreafico, M., & Bugnyar, T. (2015). With whom to dine? Ravens' responses to food-associated calls depend on individual characteristics of the caller. *Animal Behaviour*, *99*, 33-42. https://doi.org/10.1016/j.anbehav.2014.10.015

tabulate (Version 0.9.0) [Python package]. https://pypi.org/project/tabulate/

The pandas development team. (2020). *Pandas-dev/pandas:* Pandas (Version 2.2.3) [Python pacakage]. Zenodo. https://doi.org/10.5281/zenodo.3509134

Thottakara, R. (2022). *Interpersonelle Synchronität bei Kindergartenkindern: Die Rolle von Freundschaft* [Interpersonal synchrony in pre-school children: The role of friendship] (Master’s thesis, University of Vienna, Faculty of Psychology). University Library of Vienna. https://utheses.univie.ac.at/detail/64481/

Turgeon, S., & Lanovaz, M. J. (2020). Tutorial: Applying Machine Learning in Behavioral Research. *Perspectives on behavior science, 43*(4), 697–723. https://doi.org/10.1007/s40614-020-00270-y

Tyack, P. L. (2008). Convergence of calls as animals form social bonds, active compensation for noisy communication channels, and the evolution of vocal learning in mammals. *Journal of Comparative Psychology, 122*(3), 319–331. https://doi.org/10.1037/a0013087

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … van Mulbregt, P. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), 261–272. https://doi.org/10.1038/s41592-019-0686-2

Wentzel, K. R., Jablansky, S., & Scalise, N. R. (2018). Do friendships afford academic benefits? A meta-analytic study. *Educational Psychology Review, 30*(4), 1241–1267. https://doi.org/10.1007/s10648-018-9447-5

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer. https://ggplot2.tidyverse.org

Wickham, H. (2023). *Stringr: Simple, consistent wrappers for common string operations* (Version 1.5.0) [R package]. https://CRAN.R-project.org/package=stringr

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *Dplyr: A grammar of data manipulation* (Version 1.1.4) [R package]. https://CRAN.R-project.org/package=dplyr

Wickham, H., Vaughan, D., & Girlich, M. (2024). *Tidyr: Tidy messy data* (Version 1.3.1) [R package]. https://CRAN.R-project.org/package=tidyr

Xu, T. L., de Barbaro, K., Abney, D. H., & Cox, R. F. A. (2020). Finding Structure in Time: Visualizing and Analyzing Behavioral Time Series. *Frontiers in Psychology, 11*. https://www.frontiersin.org/articles/10.3389/fpsyg.2020.01457

Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on psychological science: A journal of the Association for psychological science, 12*(6), 1100–1122. https://doi.org/10.1177/1745691617693393

## Appendix A

### Compensation for radial-pull distortion in periphery

| Video observation | Formula derived for x-coordinates | Formula derived for y-coordinates |
|---|---|---|
| DYAD02F | $x_t = 0.8639 * x – 0.0154 * y$ | $y_t = - 0.0023 * x + 0.9953 * y$ |
| DYAD06F | $x_t = 0.5881 * x – 0.0204 * y$ | $y_t = - 0.0028 * x + 0.6999 * y$ |
| DYAD06NF | $x_t = 0.6016 * x – 0.0021 * y$ | $y_t = - 0.0041 * x + 0.7009 * y$ |
| DYAD10F | $x_t = 0.7280 * x – 0.0025 * y$ | $y_t = - 0.0054 * x + 0.8282 * y$ |
| DYAD10NF | $x_t = 0.7069 * x – 0.0326 * y$ | $y_t = - 0.0090 * x + 0.8199 * y$ |
| DYAD11F | $x_t = 0.6960 * x + 0.0230 * y$ | $y_t = - 0.0025 * x + 0.7516 * y$ |
| DYAD11NF | $x_t = 0.7118 * x – 0.0157 * y$ | $y_t = - 0.0023 * x + 0.7550 * y$ |
| DYAD12F | $x_t = 0.7494 * x + 0.0299 * y$ | $y_t = - 0.0149 * x + 0.8174 * y$ |
| DYAD12NF | $x_t = 0.7164 * x + 0.0052 * y$ | $y_t = - 0.0025 * x + 0.7733 * y$ |
| DYAD14F | $x_t = 0.7392 * x – 0.0106 * y$ | $y_t = - 0.0033 * x + 0.7882 * y$ |
| DYAD14NF | $x_t = 0.7178 * x – 0.0145 * y$ | $y_t = - 0.0100 * x + 0.8203 * y$ |
| DYAD15F | $x_t = 0.7196 * x – 0.0182 * y$ | $y_t = - 0.0067 * x + 0.8424 * y$ |
| DYAD15NF | $x_t = 0.7285 * x – 0.0011 * y$ | $y_t = - 0.0021 * x + 0.7762 * y$ |
| DYAD16F | $x_t = 0.7278 * x – 0.0051 * y$ | $y_t = - 0.0052 * x + 0.7955 * y$ |
| DYAD16NF | $x_t = 0.7007 * x – 0.0099 * y$ | $y_t = - 0.0035 * x + 0.7777 * y$ |
| DYAD18F | $x_t = 0.7291 * x + 0.0020 * y$ | $y_t = - 0.0056 * x + 0.8502 * y$ |
| DYAD18NF | $x_t = 0.7207 * x – 0.0054 * y$ | $y_t = - 0.0060 * x + 0.7985 * y$ |
| DYAD21F | $x_t = 0.7206 * x + 0.0078 * y$ | $y_t = - 0.0057 * x + 0.8253 * y$ |
| DYAD21NF | $x_t = 0.7169 * x – 0.0270 * y$ | $y_t = - 0.0064 * x + 0.8332 * y$ |
| DYAD23F | $x_t = 0.7407 * x – 0.0232 * y$ | $y_t = - 0.0021 * x + 0.7704 * y$ |
| DYAD23NF | $x_t = 0.7179 * x + 0.0203 * y$ | $y_t = - 0.0023 * x + 0.7682 * y$ |
| DYAD24F | $x_t = 0.7866 * x + 0.0194 * y$ | $y_t = - 0.0129 * x + 0.8342 * y$ |
| DYAD24NF | $x_t = 0.6972 * x + 0.0129 * y$ | $y_t = - 0.0100 * x + 0.7864 * y$ |

This table lists linear coefficients (intercept and slope) derived from differences of nodes between two grids. Coefficients were estimated per video observation and applied to Cartesian coordiantes as follows: $x_t = a * x + b * y$ and $y_t = c * x + d * y$, where *a* and *b* are the linear coefficients for transformation of the x-coordinates, and *c* and *d* for the y-coordinates. The variables *x* and *y* represent the original coordinates before transformation, while the variables $x_t$ and $y_t$ represent the corresponding transformed coordinates.

# Appendix B

## Testing of assumptions for linear mixed-effect models

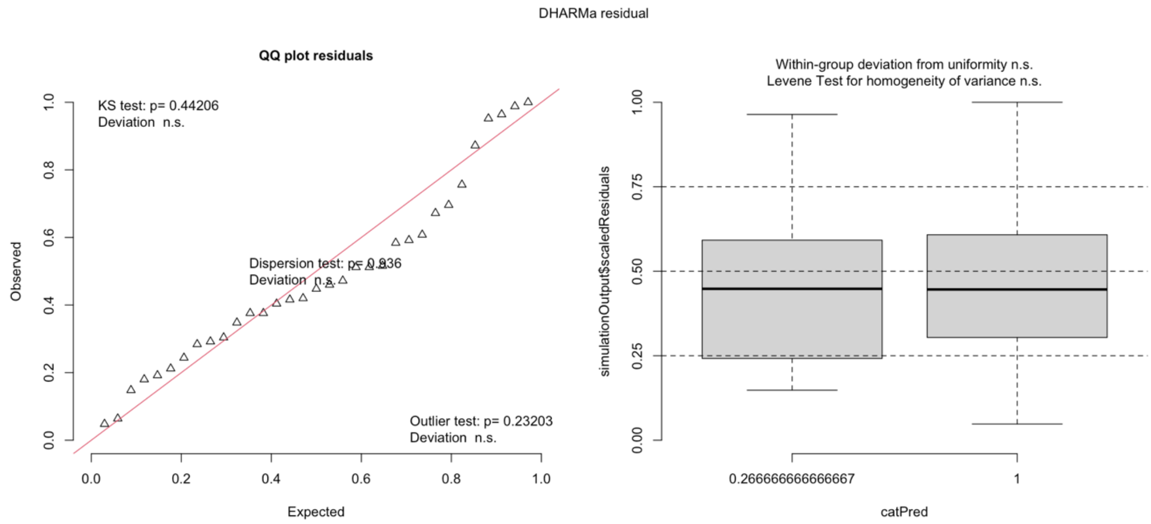



**Figure B1. DHARMa residual diagnostics for the mixed-effects model with random intercepts and dyadic distance as outcome variable.** Left: Quantile–quantile (QQ) plot of scaled residuals. The Kolmogorov–Smirnov test ($p$ = .44), dispersion test ($p$ = .84), and outlier test ($p$ = .23) did not reveal significant deviation from expected residual distribution. Right: Boxplot of residuals by predictor level showing no evidence of heteroscedasticity (Levene's test: n.s.).

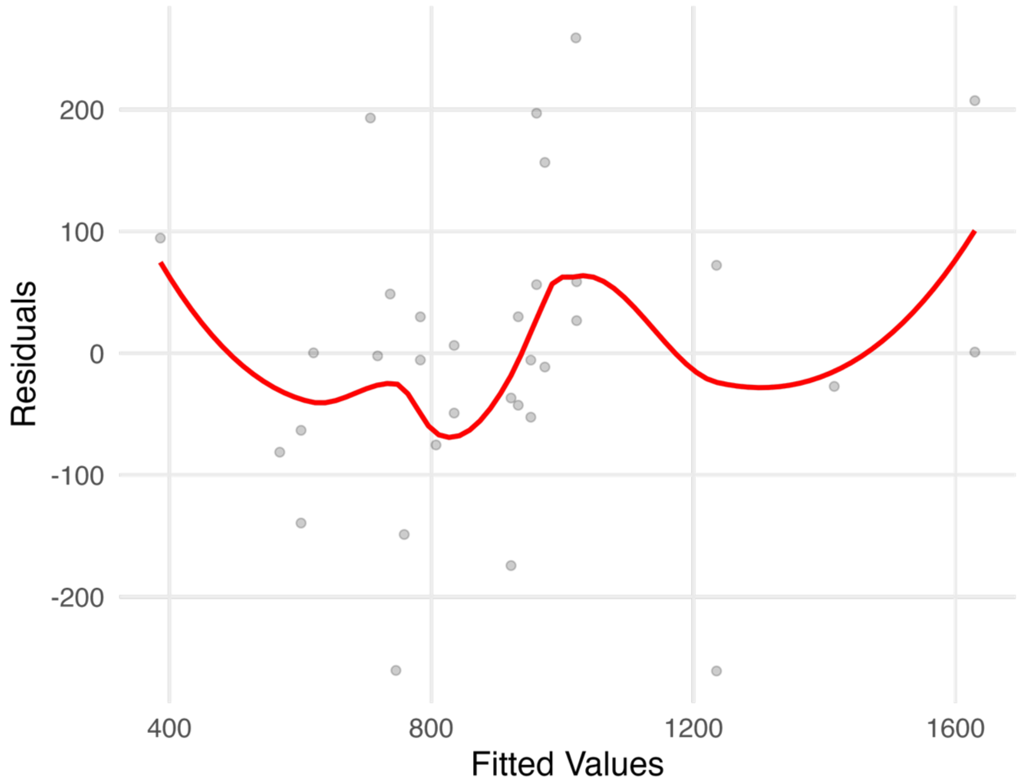



**Figure B2. Residuals vs. fitted values plot for the mixed-effects model with random intercepts and dyadic distance as outcome variable.** Even though LOESS curve (red) indicates mild non-linearity, residuals appear randomly scattered around zero, with no strong patterns or funnel shape.

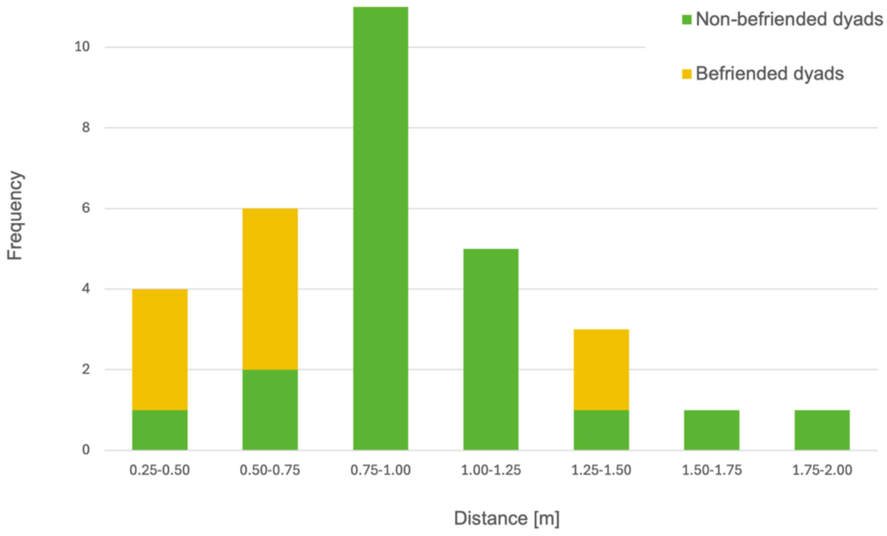



**Figure B3. Histogram of averaged dyadic distances grouped by friendship.** Distribution of averaged dyadic distances for befriended (yellow) and non-befriended (green) dyads.

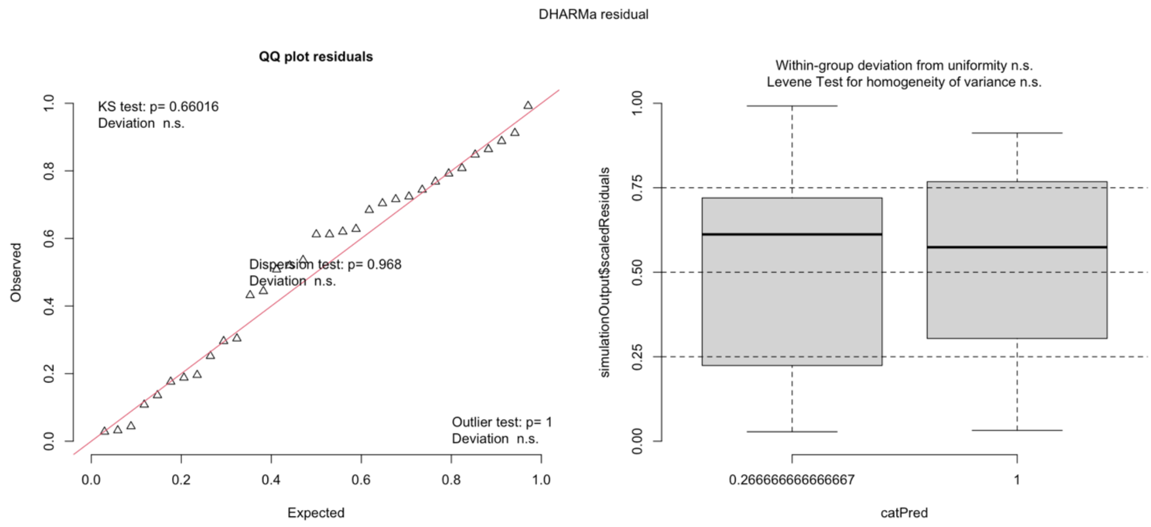



**Figure B4. DHARMa residual diagnostics for the mixed-effects model with random intercepts and dyadic social orientation angle as outcome variable.** Left: Quantile–quantile (QQ) plot of scaled residuals. The Kolmogorov–Smirnov test ($p$ = .66), dispersion test ($p$ = .97), and outlier test ($p$ = 1) did not reveal significant deviation from expected residual distribution. Right: Boxplot of residuals by predictor level showing no evidence of heteroscedasticity (Levene's test: n.s.).

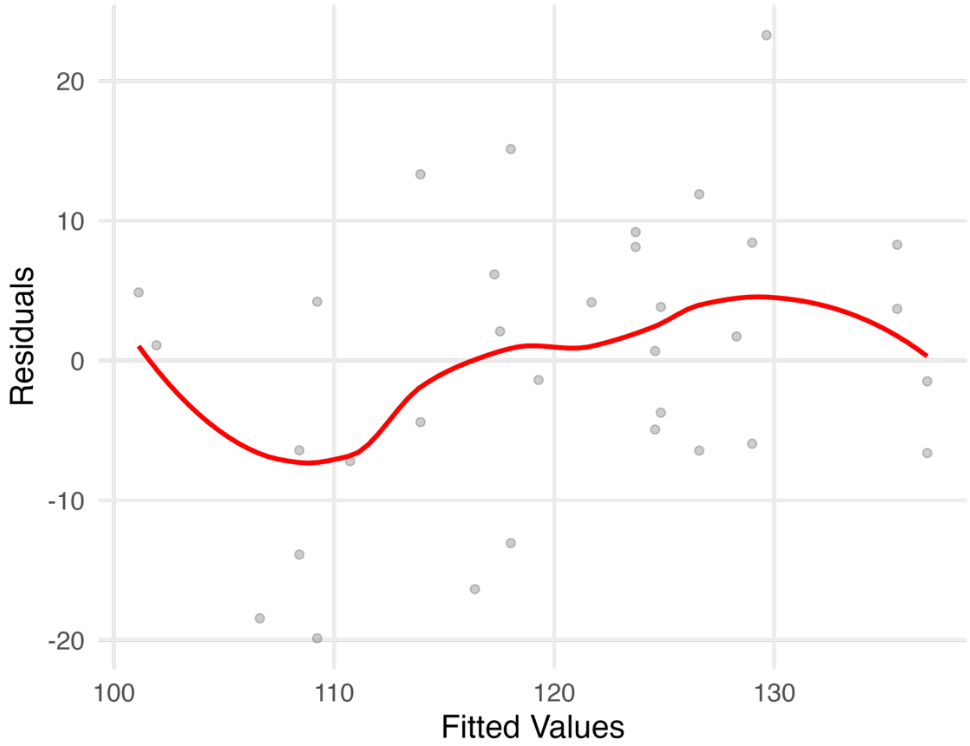



**Figure B5. Residuals vs. fitted values plot for the mixed-effects model with random intercepts and dyadic social orientation angle as outcome variable.** Residuals are randomly scattered around zero, and the LOESS curve (red) shows only mild non-linearity, indicating no strong violation of model assumptions.

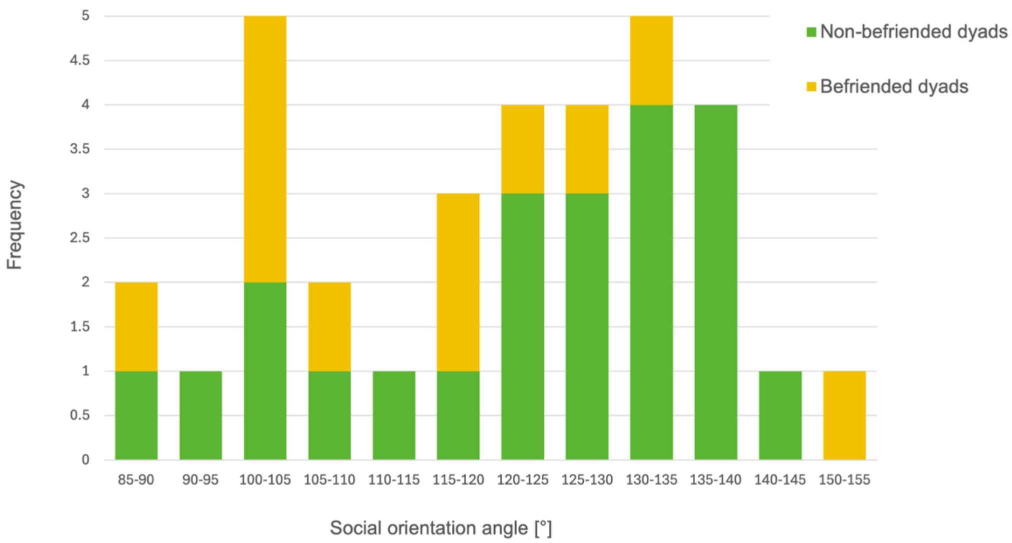
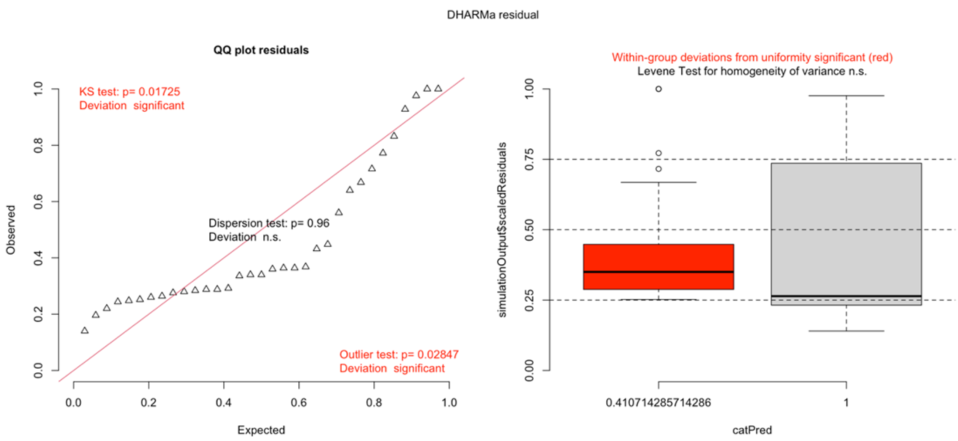



**Figure B6. Histogram of averaged dyadic social orientation angles grouped by friendship.**

Distribution of averaged dyadic social orientation angles for befriended (yellow) and non-befriended (green) dyads.

(A) *DHARMa* diagnostics of model with original outcome




(B) *DHARMa* diagnostics of model with log-transformed outcome

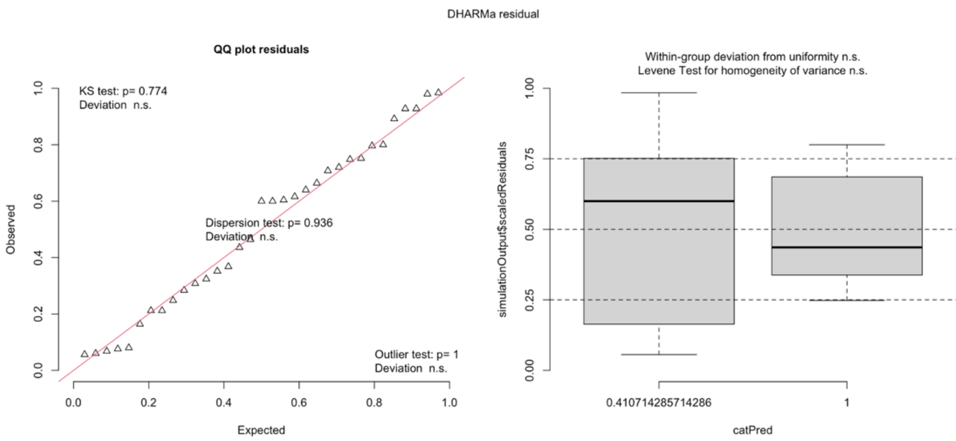



**Figure B7. DHARMa residual diagnostics for the mixed-effects model with random intercepts and slopes with durations of clear interactions as outcome variable, before and after log-transformation.** Panel (A): Left: Quantile–quantile (QQ) plot of scaled residuals. While the dispersion test was non-significant ($p$ = .96), suggesting no overdispersion, the Kolmogorov-Smirnov test ($p$ = .017) and outlier test ($p$ = .029) both indicated significant deviations from uniformity. Right: The test for within-group deviations was significant (highlighted in red), although Levene's test for homogeneity of variance was non-significant, indicating no heteroscedasticity. Panel (B): The Kolmogorov–Smirnov test ($p$ = .77), dispersion test ($p$ = .94), and outlier test ($p$ = 1) did not reveal significant deviation from expected residual distribution. Right: Boxplot of residuals by predictor level showing no evidence of heteroscedasticity (Levene's test: n.s.).

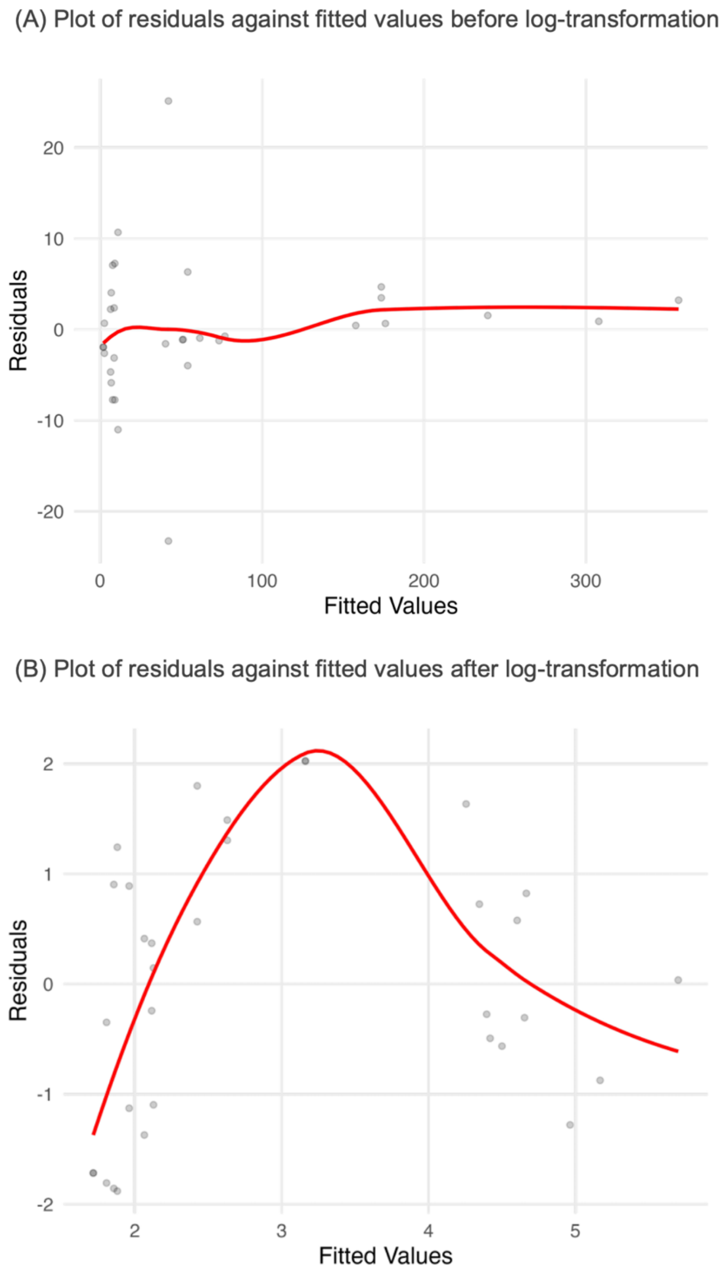



**Figure B8. Residuals vs. fitted values plot for the mixed-effects model with random intercepts and slopes with durations of clear interactions as outcome variable, before and after log-transformation.** Panel (A): Residuals show decreasing spread with larger fitted values, indicating heteroscedasticity. Panel (B): Residuals show no fan-shaped pattern, suggesting the log-transformation corrected for heteroscedasticity. However, the LOESS line (red) indicates non-linear distribution of residuals.

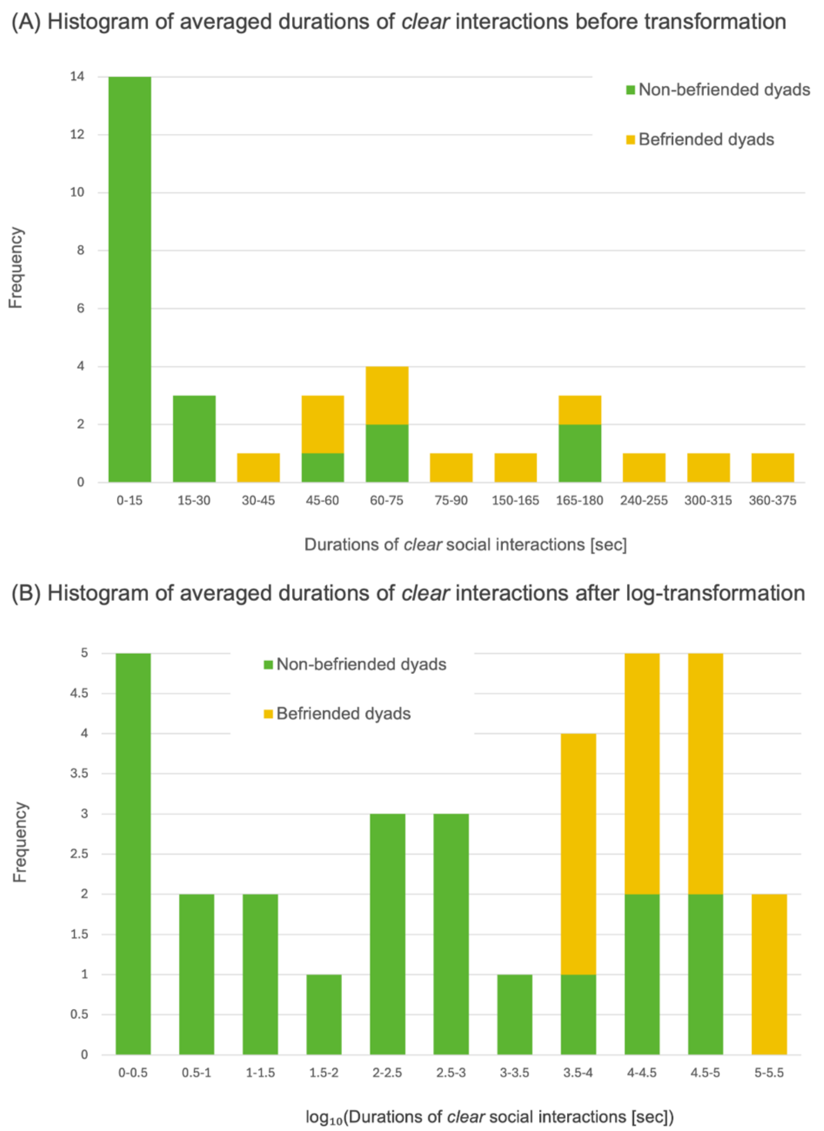
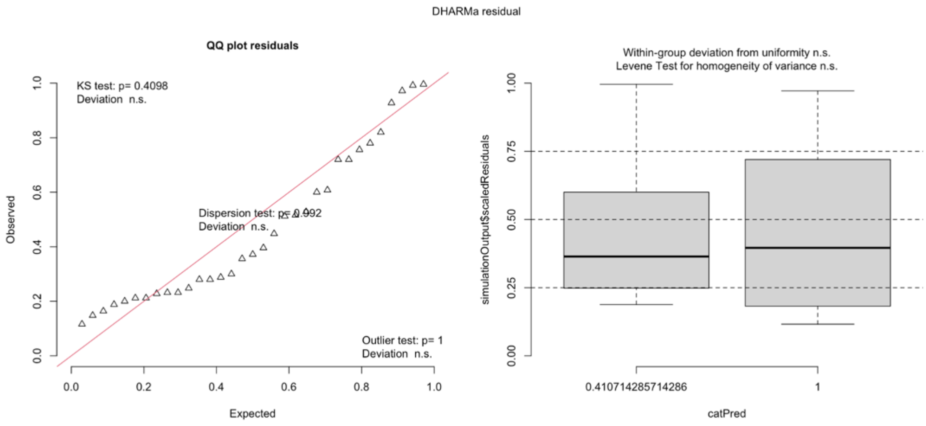



**Figure B9. Distribution of averaged durations of clear social interactions grouped by friendship, before and after transformation.** Histograms of durations of *clear* social interactions grouped for befriended (yellow) and non-befriended (green) dyads. Panel (A) shows right-skewed distribution before transformation. Panel (B) shows reduced skewedness after log-transformation.

(A) *DHARMa* diagnostics of model with original outcome




(B) *DHARMa* diagnostics of model with log-transformed outcome

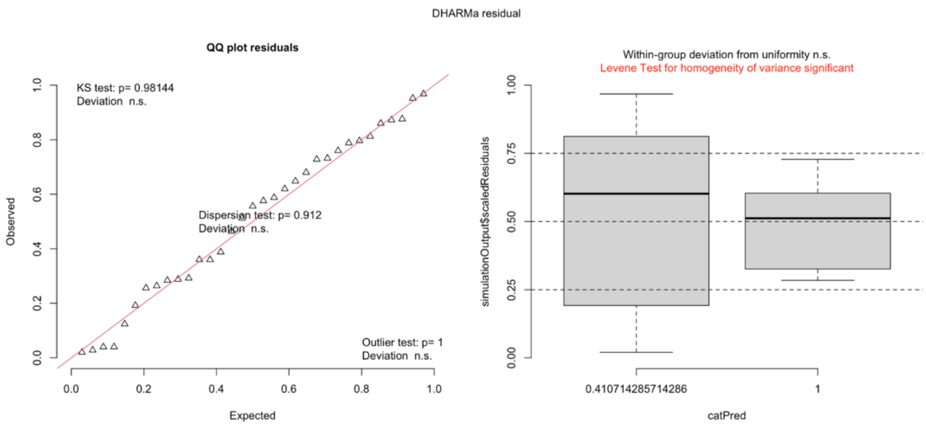



**Figure B10. DHARMa residual diagnostics for the mixed-effects model with durations of clear and unclear interactions combined as outcome variable, before and after log-transformation.** Panel (A): Left: Quantile–quantile (QQ) plot of scaled residuals. The Kolmogorov–Smirnov test ($p$ = .410), dispersion test ($p$ = .99), and outlier test ($p$ = 1) did not reveal significant deviation from expected residual distribution. Right: Boxplot of residuals by predictor level showing no evidence of heteroscedasticity (Levene's test: n.s.). Panel (B): After log-transformation of the outcome, the Kolmogorov–Smirnov test ($p$ = .98), dispersion test ($p$ = .91), and outlier test ($p$ = 1) remained non-significant. Right: However, the Levene's test for homogeneity of variance was significant.

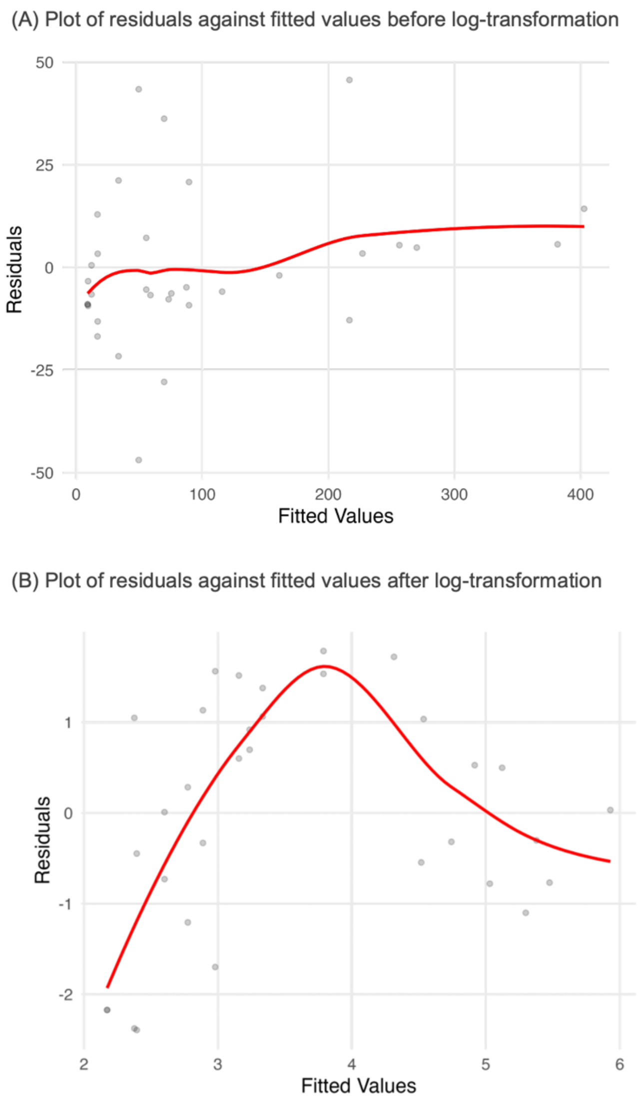



**Figure B11. Residuals vs. fitted values plot for the mixed-effects model with durations of clear and unclear interactions combined as outcome variable, before and after log-transformation.**

Panel (A): Residuals show decreasing spread with larger fitted values, indicating heteroscedasticity. Panel (B): The distribution of residuals somewhat improved regarding homoscedasticity, suggesting the log-transformation was effective. However, the LOESS line (red) indicates non-linear distribution of residuals.

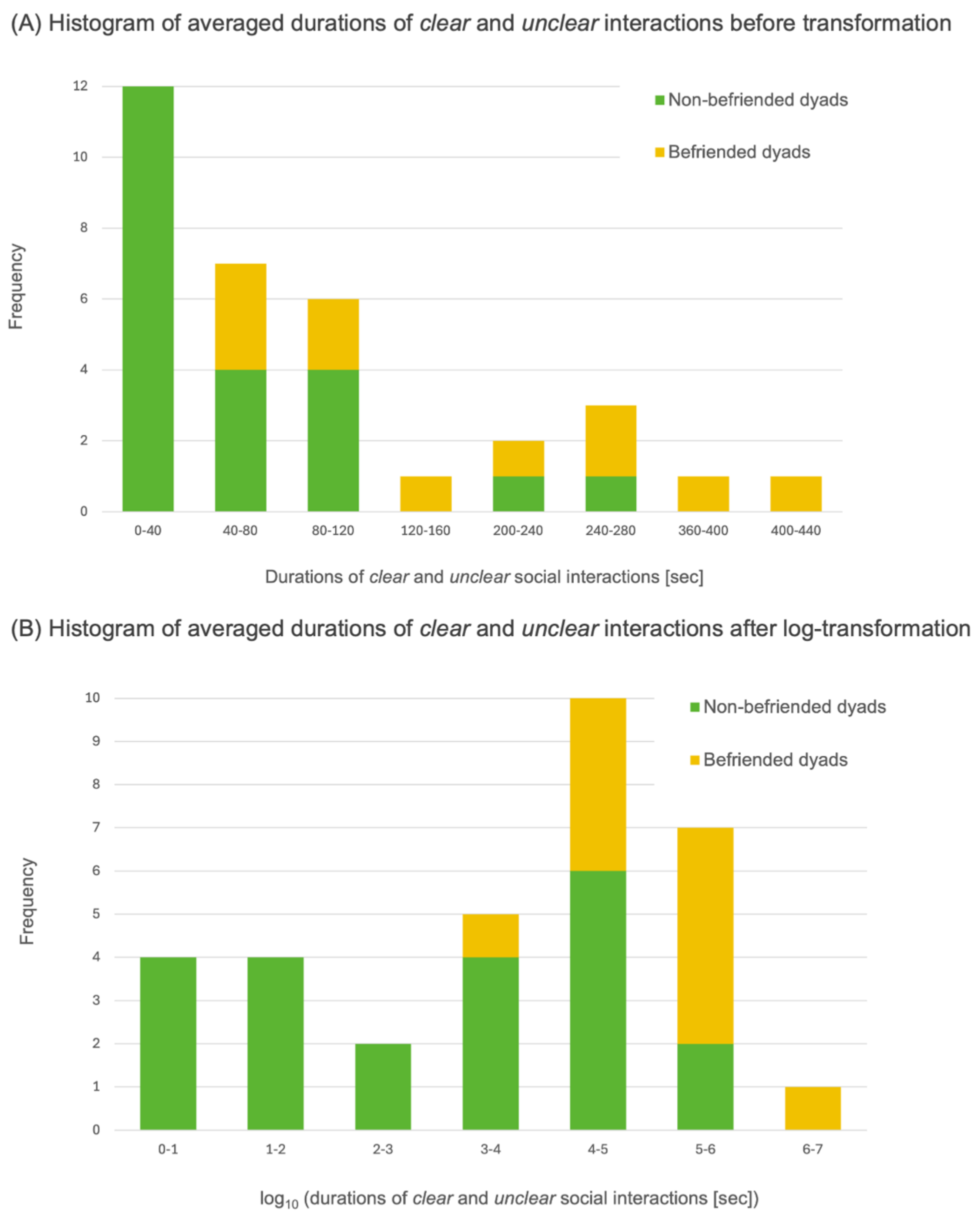
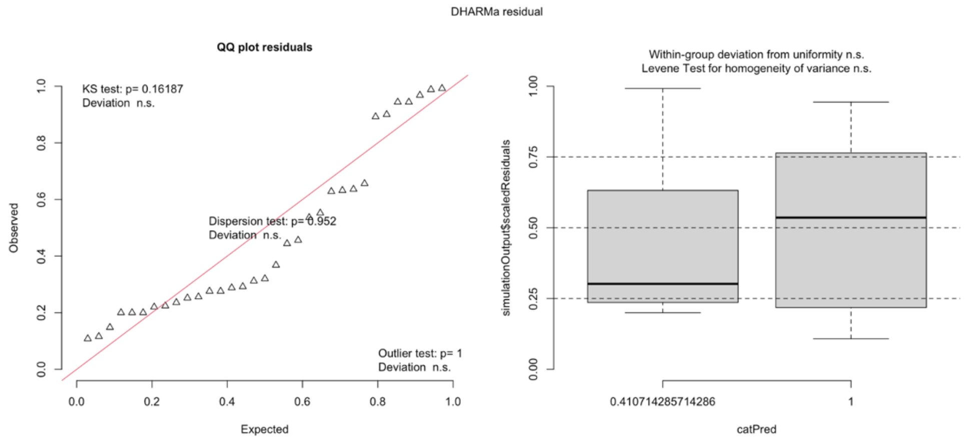



**Figure B12. Distribution of averaged durations of clear and unclear social interactions combined, grouped by friendship, before and after transformation.** Histograms of durations of *clear* social interactions grouped for befriended (yellow) and non-befriended (green) dyads. Panel (A) shows right-skewed distribution before transformation. Panel (B) shows reduced skewedness after log-transformation.

(A) *DHARMa* diagnostics of model with original outcome




(B) *DHARMa* diagnostics of model with log-transformed outcome

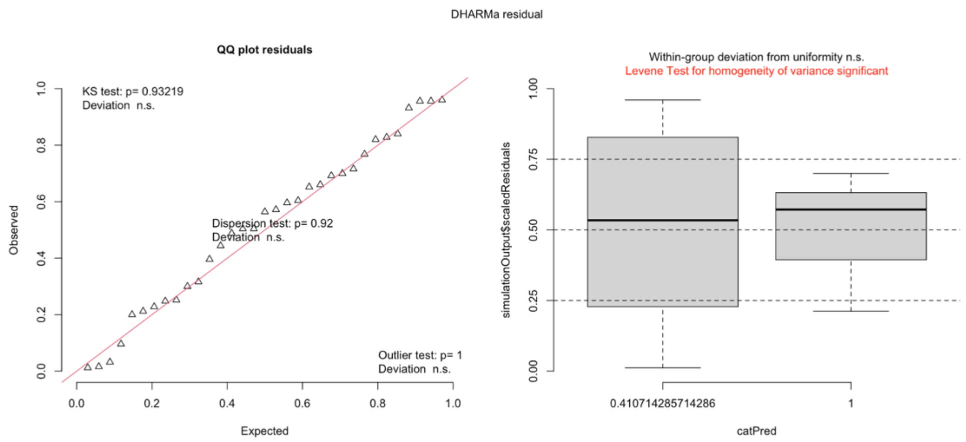



**Figure B13. DHARMa residual diagnostics for the mixed-effects with durations of all types of interactions combined as outcome variable, before and after log-transformation.** Panel (A): Left: Quantile–quantile (QQ) plot of scaled residuals. The Kolmogorov–Smirnov test ($p$ = .162), dispersion test ($p$ = .952), and outlier test ($p$ = 1) did not reveal significant deviation from expected residual distribution. Right: Boxplot of residuals by predictor level showing no evidence of heteroscedasticity (Levene's test: n.s.). Panel (B): After log-transformation of the outcome, the Kolmogorov–Smirnov test ($p$ = .932), dispersion test ($p$ = .920), and outlier test ($p$ = 1) remained non-significant. Right: However, the Levene's test for homogeneity of variance was significant.

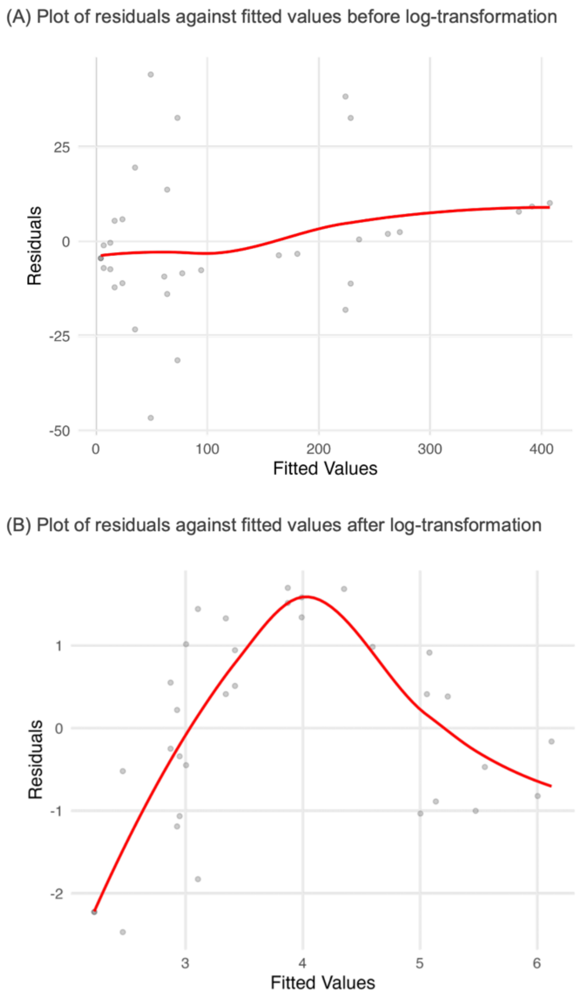



**Figure B14. Residuals vs. fitted values plot for the mixed-effects model with durations of all types of social interactions combined as outcome variable, before and after log-transformation.** Panel (A): Residuals show decreasing spread with larger fitted values, indicating heteroscedasticity. Panel (B): The distribution of residuals somewhat improved regarding homoscedasticity, suggesting the log-transformation was effective. However, the LOESS line (red) indicates non-linear distribution of residuals.

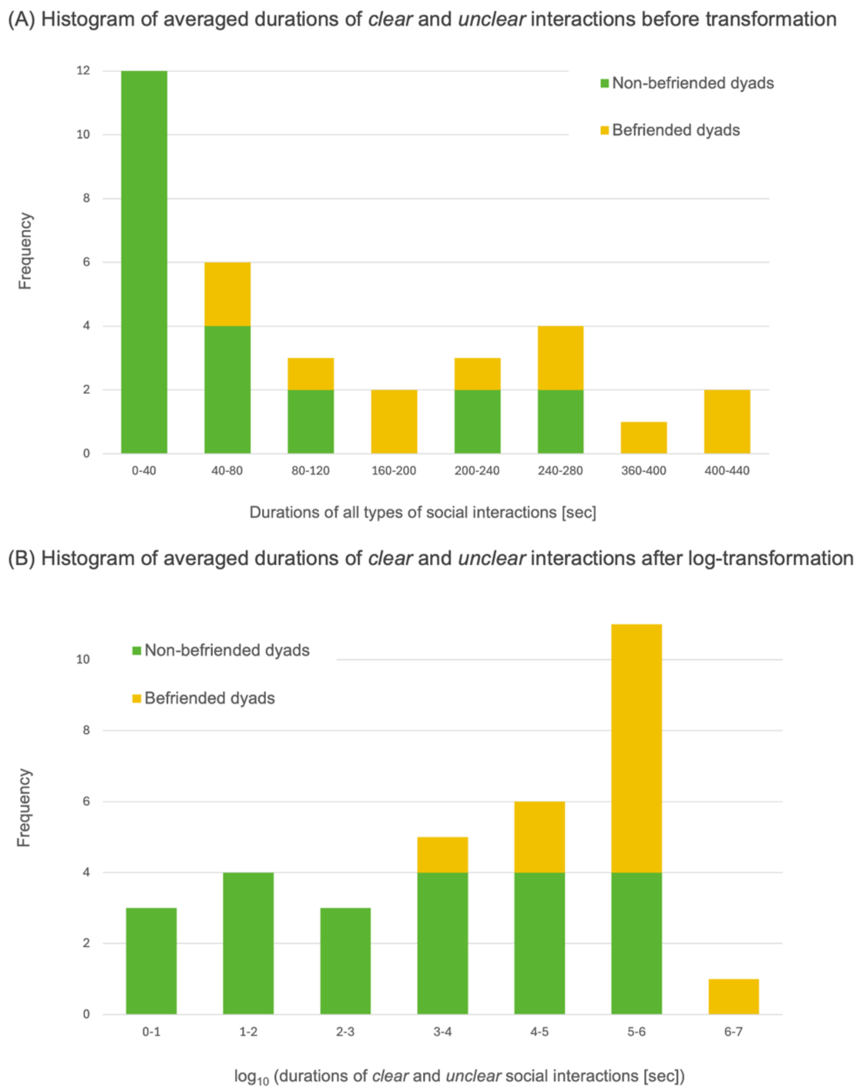



**Figure B15. Distribution of averaged durations of all types of social interactions combined, grouped by friendship, before and after transformation***.* Histograms of durations of *clear* social interactions grouped for befriended (yellow) and non-befriended (green) dyads. Panel (A) shows right-skewed distribution before transformation. Panel (B) shows histogram after log-transformation.

## Appendix C

## Performance and averaged estimates of machine learning models

(A) Training performance

| | | Actual | |
|---|---|---|---|
| | | No interaction | Interaction |
| Predicted | No interaction | 14751 | 1398 |
| | Interaction | 2861 | 16214 |
| | Sum | 17612 | 17612 |

(B) Test performance

| | | Actual | |
|---|---|---|---|
| | | No interaction | Interaction |
| Predicted | No interaction | 5119 | 100 |
| | Interaction | 393 | 279 |
| | Total | 5512 | 379 |

**Figure C1. Exemplar confusion matrix for the additive model including distance and angle predictors tested on DYAD06NF predicting clear social interactions.** Confusion matrices display the predicted and actual class labels tested on DYAD06NF, using a mixed-effects logistic regression model with random slopes including distance, angle and constellation as predictors. Panel (A) illustrates training performance after stratified undersampling, which ensured balanced class distribution (the number of actual interaction cases equals the number of actual non-interaction cases). Panel (B) shows the model's generalization to the test set, with high precision (0.981) and recall (0.929) in identifying *clear* social interactions. The test model correctly classified 5,119 instances of class 0 and 279 instances of class 1. It misclassified 393 instances of class 0 as class 1 (false positives) and 100 instances of class 1 as class 0 (false negatives).

**Table C1. Averaged performance metrics for predicting clear interactions**

| Model | Accuracy [%] | 95 % CI | AUC | F1 | Precision | Recall |
|---|---|---|---|---|---|---|
| Distance | 81.26 | [78.13, 84.39] | 0.910 | 0.873 | 0.974 | 0.800 |
| Angle | 64.30 | [60.04, 68.55] | 0.703 | 0.748 | 0.921 | 0.640 |
| Distance and angle (additive) | 83.81 | [80.83, 86.79] | 0.928 | 0.891 | 0.977 | 0.828 |
| Distance and angle (interaction) | 83.70 | [80.69, 86.71] | 0.928 | 0.891 | 0.977 | 0.826 |

All models were fitted using mixed logistic regression with random intercepts and slopes for observation as a random effect. Values reflect average test performance from leave-one-out cross-validation across 23 observations. *Additive* indicates models without interaction; *interaction* refers to inclusion of a distance × angle interaction term. Accuracy is reported as a percentage, along with its 95% confidence interval. AUC = Area Under the receiver operating characteristic Curve; F1 = harmonic mean of precision and recall.

**Table C2. Averaged fixed-effect estimates for predicting clear interactions**

| Model | Predictor | Estimate | SD | 95 % CI | *t* | *p* |
|---|---|---|---|---|---|---|
| Distance | distance | -5.44 | 0.18 | [-5.51, -5.36] | -147.47 | < .001 |
| | constellation | 0.55 | 0.16 | [ 0.49, 0.62] | 16.79 | < .001 |
| Angle | angle | -0.88 | 0.02 | [-0.88, -0.87] | -198.55 | < .001 |
| | constellation | -0.28 | 0.30 | [-0.40, -0.16] | -4.57 | < .001 |
| Distance and angle (additive) | distance | -5.47 | 0.18 | [-5.54, -5.40] | -148.53 | < .001 |
| | angle | -1.02 | 0.03 | [-1.03, -1.01] | -189.31 | < .001 |
| | constellation | 0.30 | 0.17 | [ 0.24, 0.37] | 8.77 | < .001 |
| Distance and angle (interaction) | distance | -5.58 | 0.18 | [-5.65, -5.51] | -150.37 | < .001 |
| | angle | -1.25 | 0.05 | [-1.27, -1.23] | -129.18 | < .001 |
| | interaction | -0.29 | 0.06 | [-0.32, -0.26] | -21.76 | < .001 |
| | constellation | 0.33 | 0.16 | [ 0.26, 0.40] | 9.72 | < .001 |

All models were fitted using mixed logistic regression with random intercepts and slopes of observation as a random effect. Estimates represent the average fixed-effect coefficients across all test sets. *SD* refers to the variability of the coefficient across test sets. Confidence intervals indicate range within which the true average coefficient is with 95 % probability likely to fall. *Additive* indicates models without interaction; *interaction* refers to inclusion of a distance × angle interaction term.

**Table C3. Averaged performance metrics for predicting clear and unclear interactions combined**

| Model | Accuracy [%] | 95 % CI | AUC | F1 | Precision | Recall |
|---|---|---|---|---|---|---|
| Distance | 82.79 | [79.99, 85.59] | 0.908 | 0.865 | 0.948 | 0.807 |
| Angle | 56.98 | [46.95, 67.02] | 0.616 | 0.635 | 0.828 | 0.542 |
| Distance and angle (additive) | 83.24 | [80.44, 86.05] | 0.909 | 0.871 | 0.945 | 0.817 |
| Distance and angle (interaction) | 83.30 | [80.52, 86.09] | 0.909 | 0.871 | 0.945 | 0.817 |

All models were fitted using mixed logistic regression with random intercepts and slopes for observation as a random effect. Values reflect average test performance from leave-one-out cross-validation across 23 observations. *Additive* indicates models without interaction; *interaction* refers to inclusion of a distance × angle interaction term. Accuracy is reported as a percentage, along with its 95% confidence interval. AUC = Area Under the receiver operating characteristic Curve; F1 = harmonic mean of precision and recall.

**Table C4. Averaged fixed-effect estimates for predicting clear and unclear interactions combined**

| Model | Predictor | Estimate | SD | 95 % CI | *t* | *p* |
|---|---|---|---|---|---|---|
| Distance | distance | -5.05 | 0.14 | [-5.10, -4.99] | -168.15 | < .001 |
| | constellation | 1.89 | 0.16 | [ 1.83, 1.95] | 57.62 | < .001 |
| Angle | angle | -0.39 | 0.01 | [-0.39, -0.38] | -134.65 | < .001 |
| | constellation | 1.21 | 0.27 | [ 1.10, 1.32] | 21.26 | < .001 |
| Distance and angle (additive) | distance | -4.98 | 0.14 | [-5.04, -4.92] | -174.11 | < .001 |
| | angle | -0.35 | 0.03 | [-0.36, -0.34] | -64.37 | < .001 |
| | constellation | 1.93 | 0.16 | [ 1.86, 2.00] | 57.58 | < .001 |
| Distance and angle (interaction) | distance | -4.97 | 0.14 | [-5.03, -4.91] | -173.79 | < .001 |
| | angle | -0.29 | 0.04 | [-0.31, -0.28] | -37.29 | < .001 |
| | interaction | 0.10 | 0.04 | [ 0.08, 0.11] | 11.93 | < .001 |
| | constellation | 1.92 | 0.16 | [ 1.86, 1.99] | 57.43 | < .001 |

All models were fitted using mixed logistic regression with random intercepts and slopes of observation as a random effect. Estimates represent the average fixed-effect coefficients across all test sets. *SD* refers to the variability of the coefficient across test sets. Confidence intervals indicate range within which the true average coefficient is with 95 % probability likely to fall. *Additive* indicates models without interaction; *interaction* refers to inclusion of a distance × angle interaction term.

**Table C5. Averaged performance metrics for predicting all types of interactions combined**

| Model | Accuracy [%] | 95 % CI | AUC | F1 | Precision | Recall |
|---|---|---|---|---|---|---|
| Distance | 76.8 | [72.01, 81.58] | 0.839 | 0.779 | 0.888 | 0.763 |
| Angle | 52.18 | [49.63, 66.62] | 0.639 | 0.614 | 0.793 | 0.552 |
| Distance and angle (additive) | 77.63 | [73.51, 81.76] | 0.815 | 0.809 | 0.888 | 0.776 |
| Distance and angle (interaction) | 78.23 | [74.30, 82.16] | 0.839 | 0.813 | 0.897 | 0.780 |

All models were fitted using mixed logistic regression with random intercepts and slopes for observation as a random effect. Values reflect average test performance from leave-one-out cross-validation across 23 observations. *Additive* indicates models without interaction; *interaction* refers to inclusion of a distance × angle interaction term. Accuracy is reported as a percentage, along with its 95% confidence interval. AUC = Area Under the receiver operating characteristic Curve; F1 = harmonic mean of precision and recall.

**Table C6. Averaged fixed-effect estimates for predicting all types of interactions combined**

| Model | Predictor | Estimate | SD | 95 % CI | *t* | *p* |
|---|---|---|---|---|---|---|
| Distance | distance | -3.24 | 0.16 | [-3.30, -3.17] | -97.89 | < .001 |
| | constellation | 1.63 | 0.17 | [ 1.56, 1.70] | 45.50 | < .001 |
| Angle | angle | -0.53 | 0.02 | [-0.53, -0.52] | -156.49 | < .001 |
| | constellation | 1.24 | 0.21 | [ 1.16, 1.33] | 28.63 | < .001 |
| Distance and angle (additive) | distance | -3.25 | 0.16 | [-3.32, -3.19] | -98.82 | < .001 |
| | angle | -0.59 | 0.02 | [-0.60, -0.58] | -128.65 | < .001 |
| | constellation | 1.70 | 0.17 | [ 1.63, 1.77] | 49.28 | < .001 |
| Distance and angle (interaction) | distance | -3.48 | 0.15 | [-3.54, -3.42] | -109.27 | < .001 |
| | angle | -0.85 | 0.03 | [-0.87, -0.85] | -139.06 | < .001 |
| | interaction | -0.56 | 0.04 | [-0.57, -0.54] | -70.79 | < .001 |
| | constellation | 1.73 | 0.16 | [ 1.67, 1.80] | 51.57 | < .001 |

All models were fitted using mixed logistic regression with random intercepts and slopes of observation as a random effect. Estimates represent the average fixed-effect coefficients across all test sets. *SD* refers to the variability of the coefficient across test sets. Confidence intervals indicate range within which the true average coefficient is with 95 % probability likely to fall. *Additive* indicates models without interaction; *interaction* refers to inclusion of a distance × angle interaction term.

## Appendix D

## Deutsche Zusammenfassung

Seit Halls Arbeiten zur Proxemik werden räumliche Nähe und Körperorientierung zueinander als Verhaltensmerkmale sozialer Beziehungen diskutiert. Hinweise auf Affiliation in der frühen Kindheit sind jedoch schwer zu quantifizieren, da sich soziales Verhalten in diesem Alter durch einen schnellen Verlauf, mehrere Interaktionspartner und koordinierte Bewegungen auszeichnet. Während tragbare Sensoren und automatisierte Trackingverfahren bereits eine detaillierte Messung räumlicher Merkmale im naturalistischen Kontext ermöglichen, liegen bislang nur wenige Studien vor, die prädiktive Modellierung auf diese Daten anwenden. Daher habe ich untersucht, ob einfache räumliche Maße wie dyadische Distanz und Orientierungswinkel genutzt werden können, um soziale Interaktionen im Vorschulalter zu identifizieren, und wie sich Freundschaft auf diese Verhaltensweisen auswirkt. Die Ergebnisse zeigten, dass befreundete Dyaden kürzere Distanzen einhielten und längere Interaktionen aufwiesen als nicht befreundete Dyaden, während sich im sozialen Orientierungswinkel kein signifikanter Unterschied zeigte. Prädiktive logistische Regressionsmodelle bestätigten diesen Effekt, indem kürzere Distanzen die Wahrscheinlichkeit für die Vorhersage einer sozialen Interaktion erhöhten. Der dyadische Orientierungswinkel leistete einen zusätzlichen Erklärungsbeitrag. Die höchste Klassifikationsgenauigkeit (~83 %) und AUC (~0.90) erreichten Modelle, die sowohl Distanz als auch Winkel einschlossen; Modelle, die nur Distanz berücksichtigten, lieferten jedoch nahezu gleich gute Ergebnisse. Die Übereinstimmung der Ergebnisse zwischen inferenzstatistischen und prädiktiven Ansätzen unterstreicht die Bedeutung von zwischenmenschlicher Distanz als zuverlässigem Hinweis auf Affiliation und verdeutlicht das Potenzial von maschinellem Lernen bei der Detektion sozialer Interaktionen im Kindesalter. Über den wissenschaftlichen Beitrag hinaus könnte dieser Ansatz frühe pädagogische Interventionen unterstützen, indem er skalierbare und objektive Werkzeuge zur Beobachtung sozialer Teilhabe in naturalistischen Gruppensettings bereitstellt.