Variation in gaze understanding across the life span: A process-level perspective

Julia Prein1, Manuel Bohn1,2, Luke Maurits1, Annika Werwach1, & Daniel B. M. Haun1

1 Department of Comparative Cultural Psychology, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

2 Institute of Psychology, Leuphana University Lüneburg, Germany

Author note

The authors made the following contributions. Julia Prein: Conceptualization, Methodology, Software, Investigation, Formal Analysis, Writing - Original Draft Preparation, Writing - Review & Editing; Manuel Bohn: Conceptualization, Formal Analysis, Writing - Original Draft Preparation, Writing - Review & Editing; Luke Maurits: Formal Analysis, Writing - Review & Editing; Annika Werwach: Methodology, Investigation, Writing - Review & Editing; Daniel B. M. Haun: Supervision, Writing - Review & Editing.

Correspondence concerning this article should be addressed to Julia Prein, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany. E-mail: [julia\_prein@eva.mpg.de](mailto:julia_prein@eva.mpg.de)

Abstract

Abstract

*Keywords:* social cognition, individual differences, gaze cues, cognitive modeling

*Word count:* xxx

Variation in gaze understanding across the life span: A process-level perspective

# Introduction

* why do we care about developmental trajectory? ref to stat learning paper
* variation

## Why do we need gaze understanding?

How do humans learn about their environment and navigate through their social surroundings? One possibility to extract information from the environment is through following others’ focus of attention. Building a common ground is considered especially important in communicative interactions and shared activities (Tomasello, Hare, Lehmann, & Call, 2007).

## How does gaze following emerge?

Existing studies operationalize gaze following as the ability to follow another agent’s line of sight. As one of the most fundamental social-cognitive abilities, it has been extensively studied in infancy and early childhood. Infants as young as six months can attune their gaze to that of another agent (D’Entremont, Hains, & Muir, 1997). At the end of their first year of life, infants can follow gaze to locations outside their current visual field and move themselves to gain proper perceptual access (Moll & Tomasello, 2004).

While the emergence of gaze following has been well established, less is known about the developmental trajectory throughout childhood and adolescence. One possibility is that our social-cognitive ability in question is fully developed once emerged in infancy.However, many cognitive abilities develop with age (e.g., working memory, Gathercole, Pickering, Ambridge, & Wearing, 2004). Similarly, visual processing appears to improve with age. Therefore, children could potentially improve in gaze following, fine-tuning the performance of the already existing skill.

## The scope of infants’ gaze following ability

Though these studies suggest that young infants can align their visual attention to another’s line of sight, it does not necessarily include understanding the intentions of the other agent. Infants could simply attune their orientation or be attracted by others’ gaze without processing what exactly the other is seeing (cf. Butterworth & Jarrett’s ecological and geometric mechanism, Butterworth and Jarrett (1991)]. Therefore, it is crucial to study children’s intentional understanding of gaze.

Moore, Angelopoulos, and Bennett (1997) showed that 9-month-olds followed an agent’s gaze more, when it was accompanied by a dynamic head turn in comparison to a static head turn.

In a hiding game with two search locations, Povinelli, Reaux, Bierschwale, Allain, and Simon (1997) found that three-year-olds used gaze as a cue to locate the reward, while two-year-olds performed at chance level.

In a similar object choice paradigm with two containers, Behne, Carpenter, and Tomasello (2005) investigated whether infants understand the communicative intent behind pointing and gaze cues. In contrast to Povinelli et al. (1997), they found that already 14-month-olds used the agent’s cues to select an object. In conditions with absent-minded ‘cues’, infants performed around chance. This could be interpreted as infants recognizing the nature of this joint activity: namely, that the adult’s behavior was beneficial and relevant for their object choice.

### Head vs eye direction.

It is important to note that in many existing gaze conditions, the experimenter shifted their eyes and head in synchrony (e.g., Behne et al. (2005)). Instead of pointing towards gaze understanding, a critic could claim that the results can be explained by face direction alone.

A handful of studies approached this potential confound by separately manipulating head and eye movement. Brooks and Meltzoff (2002) implemented a comparison between eye and head orientation and found that 14-month-olds were sensitive to open versus closed eyes.

Investigating the ‘cooperative eye hypothesis’, Tomasello et al. (2007) implemented six conditions, in which an experimenter oriented towards the ceiling with their eyes only, head only (eyes closed), both head and eyes, or neither. They found that human infants relied more on the eye movement, while chimpanzees paid more attention to the head movement.

Importantly, the subjects were not presented with an object choice but their attention orientation was measured.

* (Raviv & Arnon, 2018)
* (Astor & Gredebäck, 2022)
* (Colombo, 2001)
* (Scaife & Bruner, 1975)
* (Itakura & Tanaka, 1998)
* (Carpenter, Nagell, & Tomasello, 1998) “Several other studies have attempted to determine more precisely the cue that infants are using when they follow the gaze direction of others, that is, whether they use adults’ head or eye orientation. In tasks comparing infants’ responses when the experimenters turned their head and eyes together to targets with their responses when the experimenters directed their eyes to the targets but their head remained facing forward, Corkum and Moore (1995), Lempers (1979), and Lempers, Flavell, and Flavell (1977) all found that only infants age 12 months and older responded correctly when eyes and head were oriented in the same direction and that infants at all ages (i.e., through 19 months) performed poorly when eye and head direction diverged” (p.10-11) object choice.
* (Silverstein, Feng, Westermann, Parise, & Twomey, 2021) for vertical plane
* (Zhang, Zhang, Zhang, Tang, & Liu, 2019)
* (Frischen, Bayliss, & Tipper, 2007)
* (Lee, Eskritt, Symons, & Muir, 1998)
* (Coelho, George, Conty, Hugueville, & Tijus, 2006)

## Aim of the current project

### Developmental trajectory, measuring & modeling individual differences.

In this study, we were interested in the developmental trajectory of gaze understanding. While we expect the younger children to be able to follow gaze, we aimed at assessing the differentiation of their social-cognitive ability. Our goal was *not* to establish the youngest age at which children understand gaze cues. Rather, we wanted to examine how that ability changes with age.

In our study, we focused on the communicative intents of gaze: we asked children to locate a target by following an agent’s gaze. While language demands were kept low, the participants had to actively respond and, therefore, make use of the presented gaze cue.

A unique contribution of this study is the richness of the data set. Methodological challenges arise when trying to compare data across ages from qualitatively and quantitatively different study tasks. We could circumvent these issues by applying the exact same task for the entire life span.

# Lifespan

* development & individual differences in gaze understanding
* verweis methods paper reliable differences kinder & adults.
* kontinuerliche, systematische variation, wodrin? => model

<https://osf.io/6yjz3>

(#tab:lifespan\_sample)

\*\*

| Age group | n | Age mean | Age range | Age SD |
| --- | --- | --- | --- | --- |
| 3.00 | 19 (7 female) | 3.62 | 3.04 - 3.99 | 0.31 |
| 4.00 | 17 (9 female) | 4.45 | 4.05 - 4.91 | 0.30 |
| 5.00 | 22 (13 female) | 5.56 | 5.08 - 5.99 | 0.31 |
| 6.00 | 24 (16 female) | 6.50 | 6.1 - 6.99 | 0.28 |
| 7.00 | 39 (20 female) | 7.48 | 7.04 - 7.95 | 0.25 |
| 8.00 | 41 (20 female) | 8.46 | 8.03 - 8.98 | 0.27 |
| 9.00 | 56 (29 female) | 9.46 | 9.01 - 9.96 | 0.28 |
| 10.00 | 35 (22 female) | 10.49 | 10.01 - 11 | 0.28 |
| 11.00 | 54 (26 female) | 11.43 | 11.01 - 11.96 | 0.28 |
| 12.00 | 43 (19 female) | 12.41 | 12.01 - 12.99 | 0.30 |
| 13.00 | 42 (19 female) | 13.50 | 13.09 - 13.99 | 0.27 |
| 14.00 | 20 (14 female) | 14.37 | 14.05 - 14.98 | 0.23 |
| 15.00 | 21 (11 female) | 15.56 | 15.05 - 15.98 | 0.30 |
| 16.00 | 19 (10 female) | 16.51 | 16.17 - 16.97 | 0.24 |
| 17.00 | 19 (10 female) | 17.53 | 17.01 - 17.95 | 0.28 |
| 18.00 | 2 (0 female) | 18.00 | 18 - 18 | 0.00 |
| 19.00 | 5 (4 female) | 19.00 | 19 - 19 | 0.00 |
| 20.00 | 40 (25 female) | 23.02 | 20 - 29 | 2.77 |
| 30.00 | 40 (21 female) | 34.42 | 30 - 39 | 3.00 |
| 40.00 | 40 (24 female) | 44.17 | 40 - 49 | 2.92 |
| 50.00 | 40 (21 female) | 54.38 | 50 - 59 | 3.04 |
| 60.00 | 40 (21 female) | 63.73 | 60 - 69 | 2.56 |
| 70.00 | 40 (20 female) | 72.75 | 70 - 79 | 2.44 |

## Participants

We collected data from a remote child, teenager and adult sample. For the remote child and teenager sample, we recruited participants via an internal database consisting of families living in Leipzig, Germany, who volunteered to participate in child development studies and indicated an interest in online studies.

The remote child and teenager sample consisted of 471 children, including 19 3-year-olds (mean = 3.62 years, SD = 0.31, range = 3.04 - 3.99, 7 girls), 17 4-year-olds (mean = 4.45 years, SD = 0.30, range = 4.05 - 4.91, 9 girls), 22 5-year-olds (mean = 5.56 years, SD = 0.31, range = 5.08 - 5.99, 13 girls), 24 6-year-olds (mean = 6.50 years, SD = 0.28, range = 6.10 - 6.99, 16 girls), 39 7-year-olds (mean = 7.48 years, SD = 0.25, range = 7.04 - 7.95, 20 girls), 41 8-year-olds (mean = 8.46 years, SD = 0.27, range = 8.03 - 8.98, 20 girls), 56 9-year-olds (mean = 9.46 years, SD = 0.28, range = 9.01 - 9.96, 29 girls), 35 10-year-olds (mean = 10.49 years, SD = 0.28, range = 10.01 - 11, 22 girls), 54 11-year-olds (mean = 11.43 years, SD = 0.28, range = 11.01 - 11.96, 26 girls), 43 12-year-olds (mean = 12.41 years, SD = 0.30, range = 12.01 - 12.99, 19 girls), 42 13-year-olds (mean = 13.50 years, SD = 0.27, range = 13.09 - 13.99, 19 girls), 20 14-year-olds (mean = 14.37 years, SD = 0.23, range = 14.05 - 14.98, 14 girls), 21 15-year-olds (mean = 15.56 years, SD = 0.30, range = 15.05 - 15.98, 11 girls), 19 16-year-olds (mean = 16.51 years, SD = 0.24, range = 16.17 - 16.97, 10 girls), 19 17-year-olds (mean = 17.53 years, SD = 0.28, range = 17.01 - 17.95, 10 girls), 2 17-year-olds (mean = 18 years, SD = 0, range = 18 - 18, 0 girls), 5 17-year-olds (mean = 19 years, SD = 0, range = 19 - 19, 4 girls).

Children and teenagers in our sample grow up in an industrialized, urban Central-European context. Information on socioeconomic status was not formally recorded, although the majority of families come from mixed, mainly mid to high socioeconomic backgrounds with high levels of parental education.

Adults were recruited via *Prolific* (Palan & Schitter, 2018). *Prolific* is an online participant recruitment service from the University of Oxford with a predominantly European and US-American subject pool. Participants consisted of 240 English-speaking adults that reported to have normal or corrected-to-normal vision. The sample included 40 20- to 29-year-olds (mean = 23.02 years, SD = 2.77, range = 20 - 29, 25 female), 40 30- to 39-year-olds (mean = 34.42 years, SD = 3, range = 30 - 39, 21 female), 40 40- to 49-year-olds (mean = 44.17 years, SD = 2.92, range = 40 - 49, 24 female), 40 50- to 59-year-olds (mean = 54.38 years, SD = 3.04, range = 50 - 59, 21 female), 40 60- to 69-year-olds (mean = 63.73 years, SD = 2.56, range = 60 - 69, 21 female), and 40 70- to 79-year-olds (mean = 72.75 years, SD = 2.44, range = 70 - 79, 20 female).

For completing the study, subjects were paid above the fixed minimum wage (on average £10.00 per hour; see Supplements for further detail).

## Materials

We used the continuous version of the TANGO (Prein, Bohn, Kalinke, & Haun, 2022). The task was presented as an interactive web application (see Figure 1; live demo [https://ccp-odc.eva.mpg.de/tango-demo/](https://ccp-odc.eva.mpg.de/tango-demo/.); source code <https://github.com/ccp-eva/tango-demo>). The TANGO showed satisfactory internal consistency and retest reliability [with reliability estimates *Pearson’s r* ranging from .7 to .8 for the continuous task version; Prein et al. (2022)].

Each trial presented an agent standing in a window, watching a balloon (*i.e.*, target) falling to the ground. The target then fell behind a hedge (continuous task version). The agent’s gaze followed the target’s trajectory: pupil and iris moved so that their center aligned with the target center. In test trials, the target flight was covered so that participants could not see where the target landed. Participants’ task was to locate the target by tracking the agent’s gaze. They could respond by touching on the screen.

Four familiarization trials ensured that participants understood the task and felt comfortable with the response format. Then, 15 test trials followed. Completing the 19 trials took approximately 5-10 minutes.

The outcome measure was imprecision, defined as the absolute difference between the target center and the x coordinate of the participant’s click. Target coordinates were randomly generated during runtime. Each target bin, as well as all agents and target colors, occurred equally often and and did not appear in more than two consecutive trials.

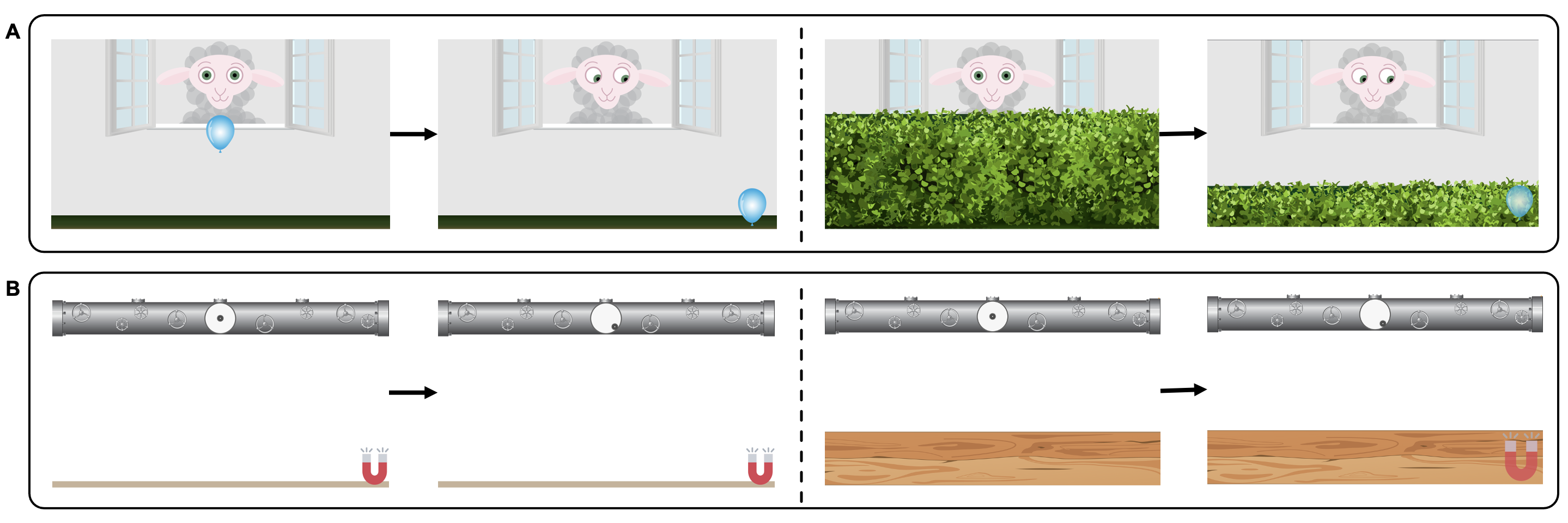


Figure1: **Setup of the TANGO and Magnet tasks**. (A) TANGO: Gaze understanding task. The agent stands in a window with the target in front of them. A hedge grows and covers the target. The target falls to a random location on the ground. The agent’s eyes track the movement of the target. (B) Magnet task: non-social vector estimation.

## Procedure

Children and teenagers received a personalized link to the study website. Caregivers were asked to provide technical support whenever needed, while explicitly being reminded to not help their children in responding. Webcam videos were recorded whenever consented and technically feasible, in order to monitor whether children and teenagers responded on their own.

## Analysis

All test trials without voice-over description were included in our analyses. We ran all analyses in R version 4.3.0 (2023-04-21) (R Core Team, 2022). Regression models were fit as Bayesian generalized linear mixed models (GLMMs) with default priors for all analyses, using the function brm from the package brms (Bürkner, 2017, 2018).

To estimate the developmental trajectory of gaze understanding and the effect of data collection mode, we fit a GLMM predicting the task performance in each trial by age (in months, z-transformed), in R: imprecision ~ age\_centered).

Imprecision was defined as the absolute click distance between the target center and the click X coordinate, scaled according to target widths, and modeled by a lognormal distribution. We inspected the posterior distribution (mean and 95% Confidence Interval (CI)) for the age and data collection estimates.

## Results

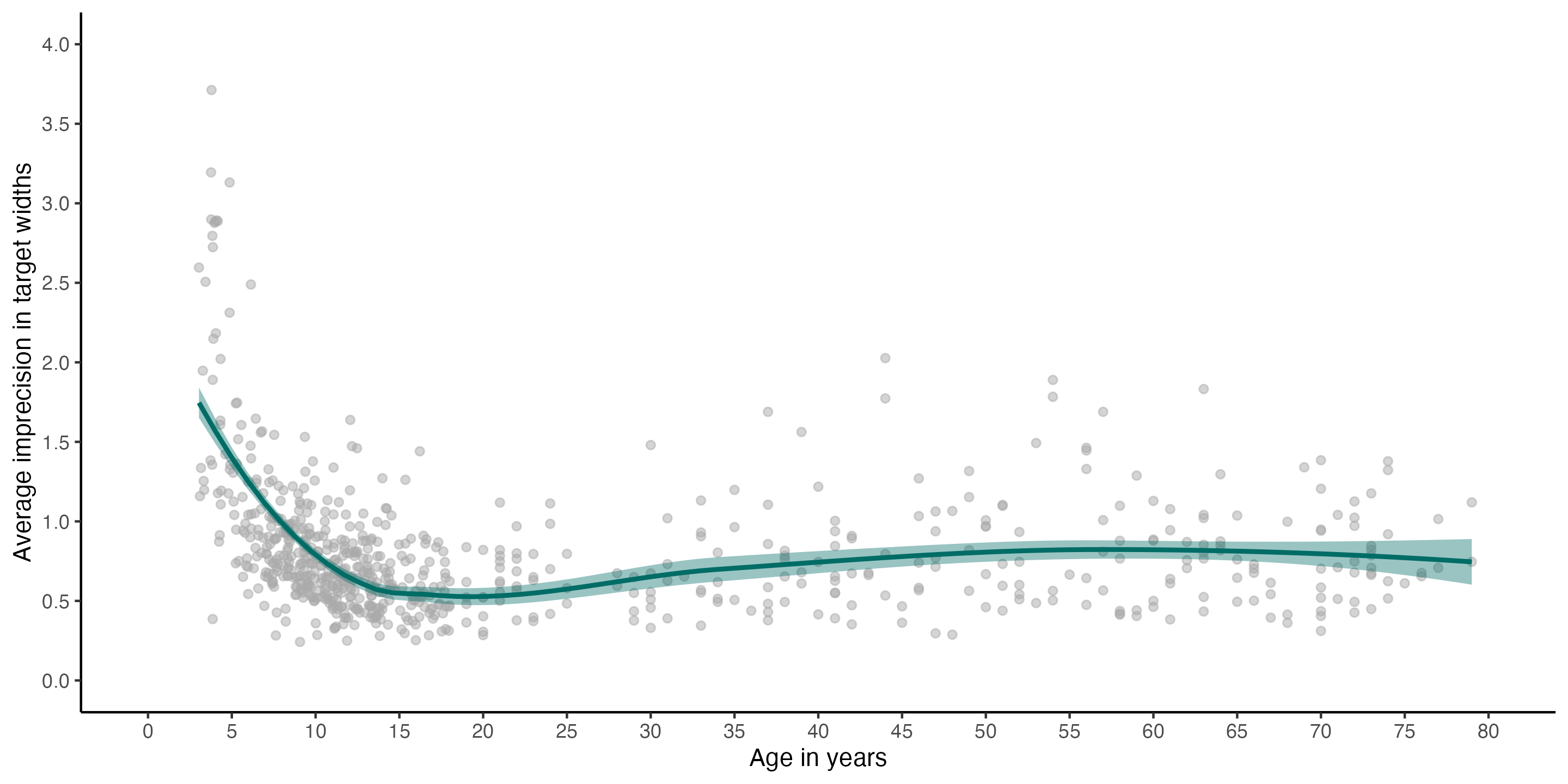


Figure2: **Differentiation in gaze understanding**. Performance is measured as imprecision, i.e., the absolute distance between the target’s center and the participant’s click (averaged across trials). The unit of imprecision is counted in the width of the target, i.e., a participant with imprecision of 1 clicked on average one target width to the left or right of the true target center.

## Discussion

Three-year-olds were surprisingly inaccurate in their responses. One possible explanation could be that they simply lacked the ability to complete the task, potentially due to issues in gaze following. Contrasting our results with previous findings on infant gaze following, this explanation is unlikely. A more likely explanation would be that children were able to follow the agent’s gaze but struggled to translate this implicit understanding into active behavior.

Another point to keep in mind is that we used subtle eye movements as cues. Many existing studies let the agents move eye and head in parallel, therefore establishing a confound with greater (head) movement. Relying exclusively on the eye movement might be trickier for children than when presented with a combined eye and head orientation.

The performance of the youngest children seems more consistent with performance demands than with a failure in gaze following.

# Computational cognitive model

Task design, data collection, and sample sizes were pre-registered: <https://osf.io/r3bhn>. The study design and procedure obtained ethical clearance by the MPG Ethics commission Munich, Germany, falling under a packaged ethics application (Appl. No. 2021\_45), and was approved by an internal ethics committee at the Max Planck Institute for Evolutionary Anthropology. The research adheres to the legal requirements of psychological research with children in Germany. Data were collected between May and August 2021.

## Participants

The sample included 60 children consisting of 20 three-year-olds (mean age = 3.47 years, SD = 0.34, range = 3.07 - 3.97, 11 girls), 20 four-year-olds (mean age = 4.61 years, SD = 0.26, range = 4.09 - 4.98, 10 girls), 20 five-year-olds (mean age = 5.66 years, SD = 0.24, range = 5.01 - 5.96, 12 girls), and 50 adults from our Lifespan study (mean age = 31.92 years, SD = 12.15, range = 18 - 63, 36 female).

Data of children was collected in kindergartens located in Leipzig, Germany. The children within each kindergarten were recruited via an internal database, where each parent priorly consented to child development studies. Adults were recruited over *Prolific* (Palan & Schitter, 2018).

Since developmental change was minimal in our adult sample (see Lifespan study) and the cognitive models were computationally heavy, we decided to only include the first 50 adults that completed the study.

## Procedure

As in the previous study, participants completed the continuous version of the TANGO (Prein et al., 2022). Children were tested in a quiet room in their kindergarten, while an experimenter guided the child through the study on a tablet.

## Analysis

## Results

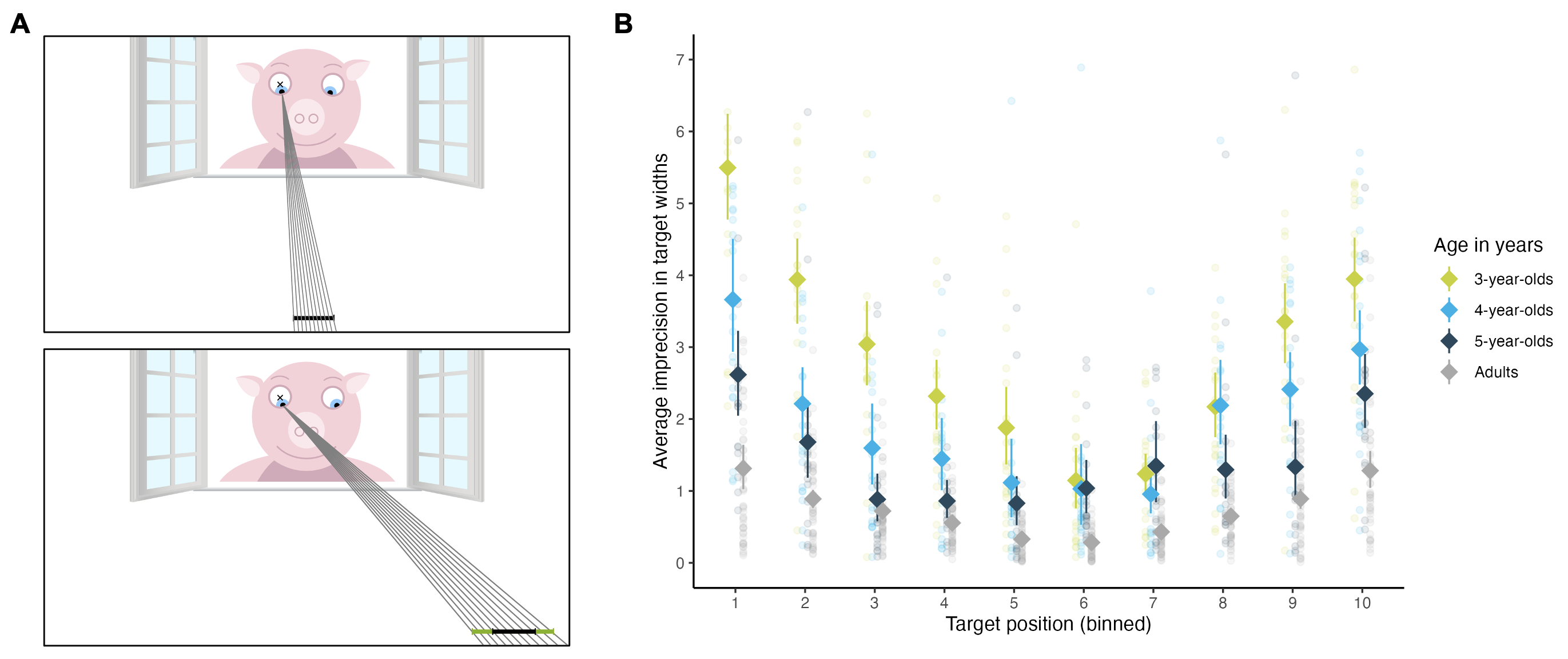


Figure3: **Gaze funnel**

## Discussion

# Components of gaze understanding

Our computational cognitive model assumes that the ability to engage in vector estimation is a crucial component of mastering gaze understanding. In this study, we sought to inquire whether there are any other cognitive processes outside vector-following that constitute gaze understanding. We aimed to assess whether there are exclusively task-specific processes at hand or whether gaze understanding recruits a general social-cognitive ability that is shared among other social-cognitive tasks.

First, we aimed at isolating the vector-following part of the gaze understanding task. For this, we designed a new non-social vector estimation task that shared all crucial design features of the understanding task. Additionally, we administered a ToM task battery, comprising 4 tasks from the ToM scale by Wellman and Liu (2004) and two additional perspective-taking tasks.

Our reasoning was that the gaze understanding task shares task demands with the non-social vector estimation task, while it shares its social context with the ToM tasks. This way, we can disentangle which components comprise gaze understanding. In this way, we can find out what contribution each component makes to explain gaze understanding.

Task design, data collection, and sample sizes were pre-registered: <https://osf.io/xsqkt>. The study design and procedure obtained ethical clearance by the MPG Ethics commission Munich, Germany, falling under a packaged ethics application (Appl. No. 2021\_45), and was approved by an internal ethics committee at the Max Planck Institute for Evolutionary Anthropology. The research adheres to the legal requirements of psychological research with children in Germany. Data were collected between February and March 2023.

## Participants

Testing took place in kindergartens in Leipzig, Germany. The sample consisted of 102 children (mean age = 4.54 years, SD = 0.31, range = 3.99 - 5.03, 54 girls).

## Procedure

Children were tested in a quiet room in their kindergarten. An experimenter guided the child through the study. Since our research questions related to individual differences and we wanted maximum control of extraneous participant variables, we employed a within-subjects design. All participants performed the following tasks in a fixed order: (1) non-social vector estimation task, (2) ToM task battery, (3) gaze understanding task. Several reasons motivated the decision. First, we decided for a fixed order to be able to compare participant’s performance straight-forwardly with each other. Second, to increase participant engagement and decrease fatigue or fuzziness, we switched from a tablet task, to tasks with personal interaction, back to a tablet task. Third, we showed the non-social vector estimation task before the gaze understanding task, so that participants would not be biased to interpret the presented stimuli as “eye- /”agent-like”.

### Non-social vector estimation.

Modeling the setup and structure of the previously applied gaze understanding task, we designed a non-social vector estimation task. This task was also presented as a webapp and made use of the concept of magnetism. The setup looked as follows. On the upper part of the screen, there was a tube with a gearwheel, located in a circular window. On the floor, there laid a magnet. The magnet then got switched on (making a cartoon-like sound), whereupon the gearwheel moved towards the magnet. The gearwheel moved in a way that its center aligned with the center of the magnet. Participants were then asked to locate the magnet. Access to the magnet’s true location was manipulated by a wooden wall: participants either had full, partial, or no visual access to the true magnet location. When no information about the magnet location was accessible, participants were expected to use the gearwheel inside the window as a non-social cue to locate the magnet.

As in the TANGO, there were three different trial types depending on the visual access to the true magnet location. In the touch training, participants had full visual access to the magnet’s location (i.e., no wooden wall). In familiarization trials, participants had partial visual access, since the wooden wall was moved in front of the target after the magnet’s location was already visible. In test trials, participants have no visual access to the magnet location, because the wall covers the magnet from the beginning of the trial.

Children received 19 trials with one touch training, two familiarization trials, and 16 test trials each. The first trial of each type comprised a voice-over description of the presented events. We conducted our analysis with 15 test trials, excluding the touch, familiarization, and voice-over trials. The outcome variable was imprecision, defined as the absolute difference between the magnet’s x coordinate and the x coordinate of the participant’s click. Magnet coordinates were generated as follows. The full width of the screen was divided into ten bins. Each bin occurred equally often, while the same bin could occur in two consecutive trials. Exact coordinates within each bin were randomly generated.

### Theory of Mind task battery.

We administered four tasks from the Wellman and Liu (2004) Theory of Mind scale. We excluded three tasks: the Diverse Desires task in order to avoid ceiling effects; and both tasks involving emotions (Belief Emotion and Real-Apparent Emotion), as we aimed at assessing the “cold, cognitive” (as compared to the “emotional”) aspects of social cognition. Instead, we added two perspective-taking level-2 tasks (Flavell, Everett, Croft, & Flavell, 1981; Flavell, Flavell, Green, & Wilcox, 1981), with the aim of increasing the task battery’s difficulty. See Supplements for further detail. The dependent variable was the aggregate score of all solved ToM tasks. In an exploratory analysis, we investigated if gaze understanding was more strongly associated with the two perspective-taking tasks compared to the other ToM tasks, as perspective-taking seems most closely theoretically related to gaze understanding (i.e., in both cases the participant is asked to judge another person’s point of view).

### Gaze understanding.

Children were tested on a tablet. As in the two previously reported studies, we presented children with the continuous version of the TANGO (Prein et al., 2022). To accentuate the social aspect of the gaze understanding task, we exchanged the animal agents (used in the previous two studies) with human faces, which were modeled after the local population in appearance. This further highlighted the contrast to the non-social vector estimation task. We had already created these stimuli for another project focusing on cross-cultural similarities in gaze understanding (<https://osf.io/tdsvc>). [[1]](#footnote-62)

## Analysis

By design, both the gaze understanding task as well as the non-social vector estimation task involve vector estimation. On the basis of the results from our computational cognitive model, we expected that children’s performance in both tasks correlate with each other. For each of these two tasks, we calculated the mean level of imprecision for each subject. We then correlated these two scores using *Pearson’s* correlation coefficients.

Regarding the relationship between the two vector estimation tasks and the ToM measures, we could imagine two possible scenarios: (A) If gaze understanding recruits a general social-cognitive ability beyond vector estimation, we expected that gaze understanding and ToM measures would correlate more strongly with each other than non-social vector estimation and ToM measures. (B) If gaze understanding relies purely on task-specific processes, then the correlation between gaze understanding and ToM measures would be comparable to the correlation between non-social vector estimation and the ToM measures. For the association between the aggregate ToM scores and the gaze understanding / non-social vector estimation tasks, we used *Spearman’s* rank correlation coefficients.

We compared the correlation between gaze understanding and ToM measures and the correlation between non-social vector estimation and ToM measures by using the Williams’ test from the function cocor.dep.groups.overlap (designed for two dependent overlapping correlations) from the package cocor (Diedenhofen & Musch, 2015).

Furthermore, to estimate which components contribute to gaze understanding, we conducted a model comparison with GLMMs predicting the mean imprecision in gaze understanding by age, non-social vector estimation, the ToM aggregate score, or the aggregate of the two perspective-taking tasks (subset of ToM battery; example of model notation in R: tango\_mean ~ age\_centered + magnet\_scaled + perspective\_scaled). The outcome variable was modeled by a lognormal distribution. With this analysis, we wanted to assess whether the ToM aggregate score or the singled-out perspective-taking score added additional explanatory value when predicting the gaze understanding score.

## Results

As expected, we found that gaze understanding as a social vector estimation task correlated with the non-social vector estimation task, *r* = 0.42, 95%CI [0.24, 0.56]. Importantly, however, the two vector estimation tasks were not redundant: only a part of the variance in gaze understanding could be explained by non-social vector estimation.

Gaze understanding and perspective-taking showed a *Spearman* correlation coefficient of = -0.27, 95%CI [-0.45, -0.08], while non-social vector estimation and perspective-taking showed a correlation of = -0.10, 95%CI [-0.29, 0.09]. According to the Williams’ test, these two correlations did not differ significantly from each other, *t*(99) = -1.63, *p* = 0.11.

Our model comparison revealed that gaze understanding was best predicted by a model including non-social vector estimation ( = 0.16, 95% CrI [0.09; 0.23]) and perspective-taking ( = -0.09; 95% CrI [-0.16, -0.03]), even when controlling for age ( = -0.11, 95% CrI [-0.34, 0.13]).

Taken together, this shows that the gaze understanding task recruited social-cognitive abilities beyond vector estimation. Evidently, it shared some of its variance with other level 2 perspective-taking tasks, while the overall ToM aggregate score did not add explanatory power.

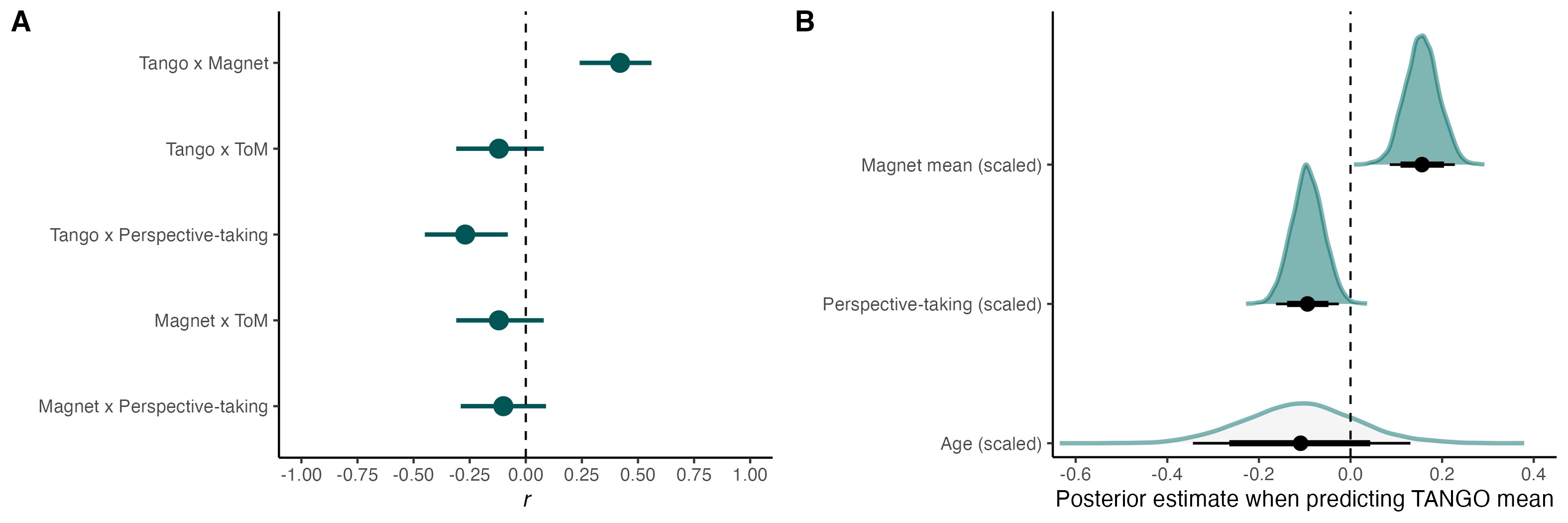


Figure4: **Components of gaze understanding**

## Discussion

In previous work, we could establish that the TANGO is suited as an individual differences measure (Prein et al., 2022). Capturing meaningful variability in performance is a crucial task feature when we are interested in revealing the relationship between different cognitive abilities.

It seems noteworthy to point out that lower correlations between ToM skills and gaze understanding could be grounded in the design features of the applied ToM tasks. The tasks we used to measure ToM skills were not designed to capture individual differences: they rely on an aggregate score of dichotomous measures. These sum scores can only capture limited variance, which may obscure potential correlations. However, since these tasks are the gold standard in the social-cognitive literature and continuous measures with satisfying psychometric properties are, to the best of our knowledge, still scarce, we nonetheless relied on them in this study. The development of new measures to capture individual differences in social-cognitive abilities like false-belief understanding seems desirable and essential to move this line of research further.

# General discussion

# Limitations

# Conclusion

# Declarations

## Open practices statement

The web application (<https://ccp-odc.eva.mpg.de/tango-demo/>) described here is open source (<https://github.com/ccp-eva/tango-demo>). The data sets generated during and/or analysed during the current study are available in the [gazecues-modeling] repository (<https://github.com/ccp-eva/gazecues-modeling>). All experiments were pre-registered (<https://osf.io/zjhsc/>).

## Funding

This study was funded by the Max Planck Society for the Advancement of Science, a noncommercial, publicly financed scientific organization (no grant number). We thank all the children, caregivers, and adults who participated in the study. We thank Jana Jurkat for her help with data collection.

## Conflicts of interest

The authors declare that they have no conflict of interest.

## Consent to participate

Informed consent was obtained from all individual participants included in the study or their legal guardians.

## Authors’ contributions

The authors made the following contributions: ### TODO

# References

Astor, K., & Gredebäck, G. (2022). Gaze following in infancy: Five big questions that the field should answer. In *Advances in Child Development and Behavior* (p. S0065240722000192). Elsevier. <https://doi.org/10.1016/bs.acdb.2022.04.003>

Behne, T., Carpenter, M., & Tomasello, M. (2005). One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science*, *8*(6), 492–499. <https://doi.org/10.1111/j.1467-7687.2005.00440.x>

Brooks, R., & Meltzoff, A. N. (2002). The importance of eyes: How infants interpret adult looking behavior. *Developmental Psychology*, *38*(6), 958–966. <https://doi.org/10.1037/0012-1649.38.6.958>

Bürkner, P.-C. (2017). Brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80*(1). <https://doi.org/10.18637/jss.v080.i01>

Bürkner, P.-C. (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal*, *10*(1), 395. <https://doi.org/10.32614/RJ-2018-017>

Butterworth, G., & Jarrett, N. (1991). What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology*, *9*(1), 55–72. <https://doi.org/10.1111/j.2044-835X.1991.tb00862.x>

Carpenter, M., Nagell, K., & Tomasello, M. (1998). [Social cognition, joint attention, and communicative competence from 9 to 15 months of age](https://www.ncbi.nlm.nih.gov/pubmed/9835078). *Monographs of the Society for Research in Child Development*, *63*(4), i–vi, 1–143.

Coelho, E., George, N., Conty, L., Hugueville, L., & Tijus, C. (2006). Searching for asymmetries in the detection of gaze contact versus averted gaze under different head views: A behavioural study. *Spatial Vision*, *19*(6), 529–545. <https://doi.org/10.1163/156856806779194026>

Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology*, *52*, 337–367. <https://doi.org/10.1146/annurev.psych.52.1.337>

D’Entremont, B., Hains, S. M. J., & Muir, D. W. (1997). A demonstration of gaze following in 3- to 6-month-olds. *Infant Behavior and Development*, *20*(4), 569–572. <https://doi.org/10.1016/S0163-6383(97)90048-5>

Diedenhofen, B., & Musch, J. (2015). Cocor: A Comprehensive Solution for the Statistical Comparison of Correlations. *PLoS ONE*, *10*(4), e0121945. <https://doi.org/10.1371/journal.pone.0121945>

Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children’s knowledge about visual perception: Further evidence for the Level 1 2 distinction. *Developmental Psychology*, *17*, 99–103. <https://doi.org/10.1037/0012-1649.17.1.99>

Flavell, J. H., Flavell, E. R., Green, F. L., & Wilcox, S. A. (1981). The Development of Three Spatial Perspective-Taking Rules. *Child Development*, *52*(1), 356–358. <https://doi.org/10.2307/1129250>

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin*, *133*(4), 694–724. <https://doi.org/10.1037/0033-2909.133.4.694>

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The Structure of Working Memory From 4 to 15 Years of Age. *Developmental Psychology*, *40*, 177–190. <https://doi.org/10.1037/0012-1649.40.2.177>

Itakura, S., & Tanaka, M. (1998). Use of experimenter-given cues during object-choice tasks by chimpanzees (Pan troglodytes), an orangutan (Pongo pygmaeus), and human infants (Homo sapiens). *Journal of Comparative Psychology*, *112*(2), 119–126. <https://doi.org/10.1037/0735-7036.112.2.119>

Lee, K., Eskritt, M., Symons, L. A., & Muir, D. (1998). Children’s use of triadic eye gaze information for "mind reading". *Developmental Psychology*, *34*(3), 525–539. <https://doi.org/10.1037//0012-1649.34.3.525>

Moll, H., & Tomasello, M. (2004). 12- and 18-month-old infants follow gaze to spaces behind barriers. *Developmental Science*, *7*(1), F1–F9. <https://doi.org/10.1111/j.1467-7687.2004.00315.x>

Moore, C., Angelopoulos, M., & Bennett, P. (1997). The role of movement in the development of joint visual attention. *Infant Behavior and Development*, *20*(1), 83–92. <https://doi.org/10.1016/S0163-6383(97)90063-1>

Palan, S., & Schitter, C. (2018). Prolific.ac subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27. <https://doi.org/10.1016/j.jbef.2017.12.004>

Povinelli, D. J., Reaux, J. E., Bierschwale, D. T., Allain, A. D., & Simon, B. B. (1997). Exploitation of pointing as a referential gesture in young children, but not adolescent chimpanzees. *Cognitive Development*, *12*(4), 423–461. <https://doi.org/10.1016/S0885-2014(97)90017-4>

Prein, J. C., Bohn, M., Kalinke, S., & Haun, D. B. M. (2022). *TANGO: A reliable, open-source, browser-based task to assess individual differences in gaze understanding in 3 to 5-year-old children and adults*. PsyArXiv. <https://doi.org/10.31234/osf.io/vghw8>

R Core Team. (2022). *R: A language and environment for statistical computing* [Manual]. Vienna, Austria: R Foundation for Statistical Computing.

Raviv, L., & Arnon, I. (2018). The developmental trajectory of children’s auditory and visual statistical learning abilities: Modality-based differences in the effect of age. *Developmental Science*, *21*(4), e12593. <https://doi.org/10.1111/desc.12593>

Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature*, *253*(5489), 265–266. <https://doi.org/10.1038/253265a0>

Silverstein, P., Feng, J., Westermann, G., Parise, E., & Twomey, K. E. (2021). Infants Learn to Follow Gaze in Stages: Evidence Confirming a Robotic Prediction. *Open Mind*, *5*, 174–188. <https://doi.org/10.1162/opmi_a_00049>

Tomasello, M., Hare, B., Lehmann, H., & Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: The cooperative eye hypothesis. *Journal of Human Evolution*, *52*(3), 314–320. <https://doi.org/10.1016/j.jhevol.2006.10.001>

Wellman, H. M., & Liu, D. (2004). Scaling of Theory-of-Mind Tasks. *Child Development*, *75*(2), 523–541. <https://doi.org/10.1111/j.1467-8624.2004.00691.x>

Zhang, X., Zhang, Z., Zhang, Z., Tang, Y., & Liu, W. (2019). The role of the motion cue in the dynamic gaze-cueing effect: A study of the lateralized ERPs. *Neuropsychologia*, *124*, 151–160. <https://doi.org/10.1016/j.neuropsychologia.2018.12.016>

1. In an exploratory analysis, we compared children’s imprecision levels in the TANGO task with animal vs. human agents. Based on a GLMM analysis, we conclude that there was no evidence of a stable effect of stimulus choice (human vs. animal). See Supplements for further detail. [↑](#footnote-ref-62)