1   Unity in diversity: Children from 17 diverse communities process gaze in similar ways

2   Manuel Bohn[1,2], Julia Prein[1,2], . . . ···, & Daniel B. M. Haun[2]

3   [1] Institute of Psychology in Education, Leuphana University Lüneburg

4   [2] Department of Comparative Cultural Psychology, Max Planck Institute for Evolutionary

5   Anthropology

6   Author Note

7   Add complete departmental affiliations for each author here. Each new line herein

8   must be indented, like this line.

9   Enter author note here.

10   The authors made the following contributions. Manuel Bohn: Conceptualization,

11   Writing - Original Draft Preparation, Writing - Review & Editing; Julia Prein:

12   Conceptualization, Methodology, Writing - Review & Editing; . . . : Conceptualization,

13   Methodology, Writing - Review & Editing; Daniel B. M. Haun: Conceptualization, Writing

14   - Review & Editing.

15   Correspondence concerning this article should be addressed to Manuel Bohn, Postal

16   address. E-mail: my@email.com

Abstract

Understanding gaze is critical for human social interaction. In a comprehensive cross-cultural study spanning five continents and 17 distinct cultural contexts, we examined the development of gaze understanding among preschoolers using a single reliable task tailored to each cultural setting. Our data provides evidence that children worldwide process gaze in the same way. Key performance signatures of a computational model that sees gaze following as a form of social vector estimation were found in all communities. Additionally, we found a subtle yet cross-culturally consistent relationship between children's environments and their ability to understand gaze. These results highlight the cross-cultural robustness of core social cognitive skills, as well as similarities in the developmental process, suggesting that fundamental aspects of social cognition and interaction emerge and develop in comparable ways globally and hence are a bedrock of human social cognition.

Keywords: keywords

Word count: X

Unity in diversity: Children from 17 diverse communities process gaze in similar ways

Social cognition is a defining aspect of the human species (*1–3*). It is supposed to enable unique forms of communication and cooperation that underlie cumulative cultural evolution and the formation of complex societies (*4–6*). The eyes are the proverbial window to the mind and the starting point for a majority of social reasoning processes (*7*). Gaze is used to infer the focus of visual attention, which is a critical aspect of coordinated activities, including communication and cooperation (*8, 9*). The ability to follow gaze emerges early in childhood (*10–12*) and individual differences in children's gaze following ability predict later life outcomes, most notably, later communicative abilities (*13, 14*). Underlying this narrative is the wide-held assumption that gaze following is fundamental to human social cognition and, therefore, works and develops in the same way across human societies despite substantial variation in developmental contexts. This claim, however, lacks a solid empirical foundation.

We conducted a large-scale cross-cultural study on the development of gaze-following abilities to study potentially universal processing signatures and their development. Previous developmental studies focused on the onset of gaze following in infancy (*15, 16*). The 1377 participants in the study lived in 17 different communities spread over five continents (Fig. 1A, Tab. 1). The countries from which data was contributed to the study represented 46 % of the world's population. Communities covered a broad spectrum of geographical locations, social and political systems, languages, and subsistence styles (see Supplemental Materials).

We used an animated picture book tablet task in which participants were asked to locate a hidden object based on observing an agent's gaze. Children watched a balloon disappear behind a hedge. An agent followed the trajectory of the balloon with their eyes (Fig. 1B). The key dependent variable was the precision with which children located the agent's focus of attention, that is, the deviation between where the agent looked (where the
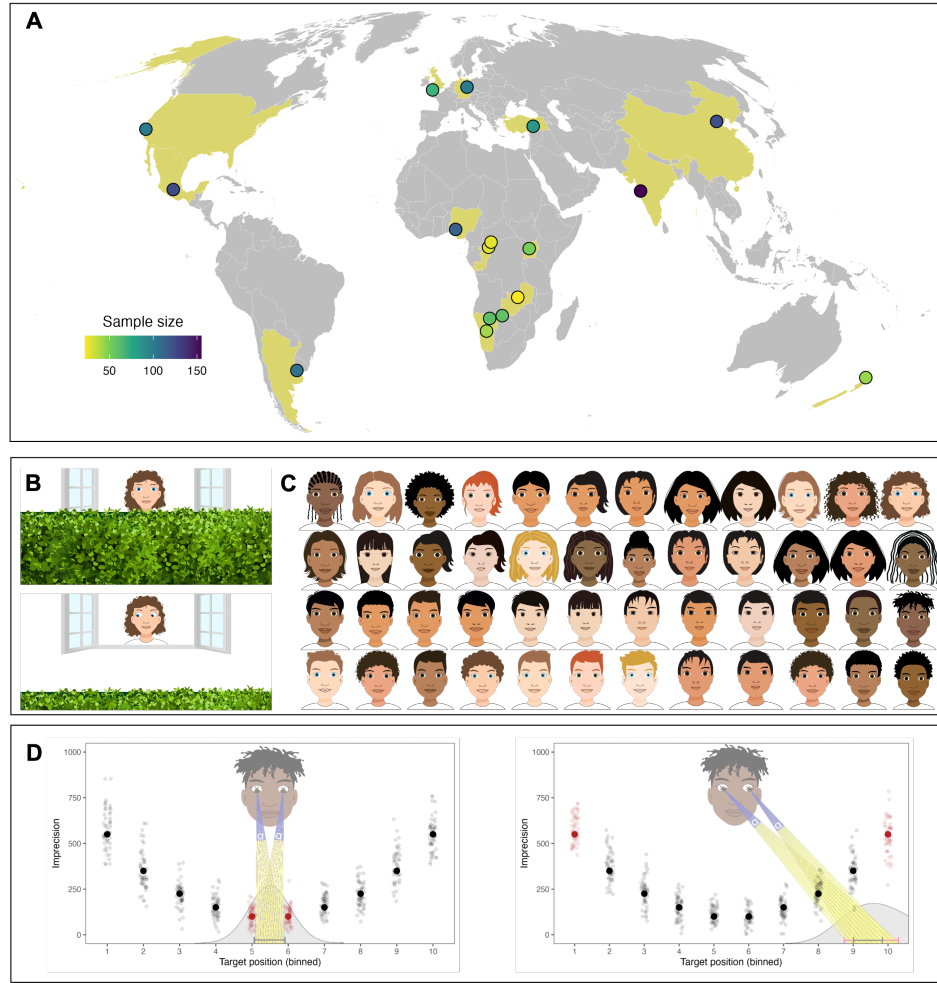
Figure 1. (A) Data collection sites. Points show approximate geographical location of the data collection sites, coloring shows the sample sizes. (B) Screenshot from the task. The scene depicts the choice phase in a test trial. Participants had to use the gaze of the agent to locate the balloon and click on the hedge where they thought the balloon was. Agents, audio recordings and backgrounds were adpated to each cultural setting. (C) Drawings used as agents across cultural settings. (D) Graphical illustration of the cognitive model. Individuals infer the target of an agent's attention by estimating a vector based on the position of the pupils within the eyes. This process is noisy, illustrated by the different vectors. Individuals differ in their level of precision (indicated by alpha). For a given level of precision, the further out the target lands, the less precise the model predicts individuals to be. Solid and transparent dots show simualted means and individual data points to illustrate the predicted effect of target position.

Table 1

Participant demographics.

Continent	Country	Community	N(male)	Age (range)	Language	Market integration	Touchscreen
America	Argentina	Buenos Aires	105 (53)	4.72 (3.00 - 6.96)	Spanish (Rioplatense)	high	0.90
	Mexico	Ocuilan	127 (63)	4.96 (2.57 - 6.95)	Spanish (Mexican)	medium	0.77
	USA	Stanford	98 (54)	4.99 (2.52 - 7.90)	English (American)	high	0.98
Africa	Namibia	Hai‖om	60 (38)	5.85 (2.74 - 8.34)	Hai‖om	low	0.05
		Khwe	59 (24)	5.84 (3.38 - 8.63)	Khwedam	low	0.19
		Windhoek	39 (17)	5.69 (2.66 - 8.66)	English (Nigerian)	high	0.95
	Nigeria	Akure	114 (54)	5.07 (2.57 - 7.33)	English (Nigerian)	high	0.91
	Rep. Congo	BaYaka	29 (13)	7.80 (3.94 - 10.56)	BaYaka	low	0.00
		Bandongo	30 (11)	7.45 (3.50 - 10.95)	Lingala	low	0.00
	Uganda	Nyabyeya	125 (62)	5.94 (2.67 - 8.92)	Bemba	medium	0.34
	Zambia	Chimfunshi	22 (5)	5.98 (2.88 - 8.00)	Swahili	medium	0.14
Europe	Germany	Leipzig	100 (48)	4.88 (2.53 - 6.95)	German	high	0.89
	UK	Plymouth	70 (30)	6.02 (2.38 - 8.94)	English (British)	high	0.99
Asia	China	Beijing	123 (62)	5.47 (2.69 - 8.48)	Mandarin	high	0.95
	India	Pune	148 (73)	6.14 (3.06 - 8.83)	English (Indian) / Marathi	high	0.93
	Türkiye	Malatya	85 (40)	5.02 (2.75 - 7.12)	Turkish	high	1.00
Oceania	New Zealand	Auckland	43 (19)	5.14 (2.81 - 8.75)	English (New Zealand)	high	0.95

58 object was) and where the child thought the agent looked. We adapted visuals and audio

59 instructions specifically to each of the 17 communities. Previous work demonstrated

60 excellent individual-level measurement properties for this task (*17*). Thus, in addition to

61 group-level trends, we were able to investigate individual-level variation.

62 As the first step, we investigated developmental gains. Across all 17 communities, we

63 found a substantial increase in average levels of precision with age (fixed effect in Bayesian

64 regression model (*18*): β = -0.30, 95% HDI (-0.40 - -0.21); range of community-level effects

65 in model: β_{min} = -0.06, 95% HDI (-0.17 - 0.06) to β_{max} = -0.59, 95% HDI (-0.70 - -0.47).

66 Nevertheless, there were also marked differences between communities (see Fig. 2A). In a

67 cross-validation procedure, in which we trained a regression model on a subset of the data

68 (training data) to later predict the held out data (testing data) (*19*), we found that a

69 model that assumed cross-cultural variation to outperform simple models in 100% of cases

70 (see Supplemental Material). At first glance, it seems that highly market-integrated

71 communities around the globe showed higher levels of precision compared to less

72 market-integrated communities (see Tab. 1). However, believe that these results are better

73 understood in terms of exposure to technology and we discuss them in more detail below.

74 Furthermore, average differences between communities were small compared to differences

75 between individuals: communities did not form homogeneous clusters but largely

76 overlapping distributions in that some individuals from communities with a lower average

77 level of precision performed better compared to some individuals from a setting with a very

78 high average level of precision. Similarly, in all communities, some 4-year-olds

79 outperformed children two years older than them (see Fig. 2A).

80   Consistent developmental gains alone cannot inform us about the cognitive processes

81 children use when locating the agent's focus of attention. Recent computational work

82 modeled gaze following as social vector estimation (*20*). When observing the eyes,

83 onlookers estimate a vector from the center of the eye through the pupil. The focus of

84 attention is the location wherever the estimated vectors from both eyes hit a surface (Fig.

85 1C). The estimation process is not perfect, but each individual has a systematic level of

86 uncertainty, which is conceptualized as the cause of individual differences. Importantly,

87 this process model predicts a clear performance signature in our cross-cultural gaze

88 following task: Trials in which the agent looks further away from the center (i.e. to the left

89 or right side of the screen) should result in lower levels of precision, compared to trials in

90 which the agent looks closer to the middle. This prediction is best understood by

91 considering a similar phenomenon: pointing a torch light to a flat surface. The width of the

92 light beam represents each individual's level of uncertainty in vector estimation. When the

93 torch is directed straight down, the light beam is concentrated in a relatively small area.

94 When the torch is rotated to the side, the light from one half of the cone must travel

95 further than the light from the other half to reach the surface. As a consequence, the light
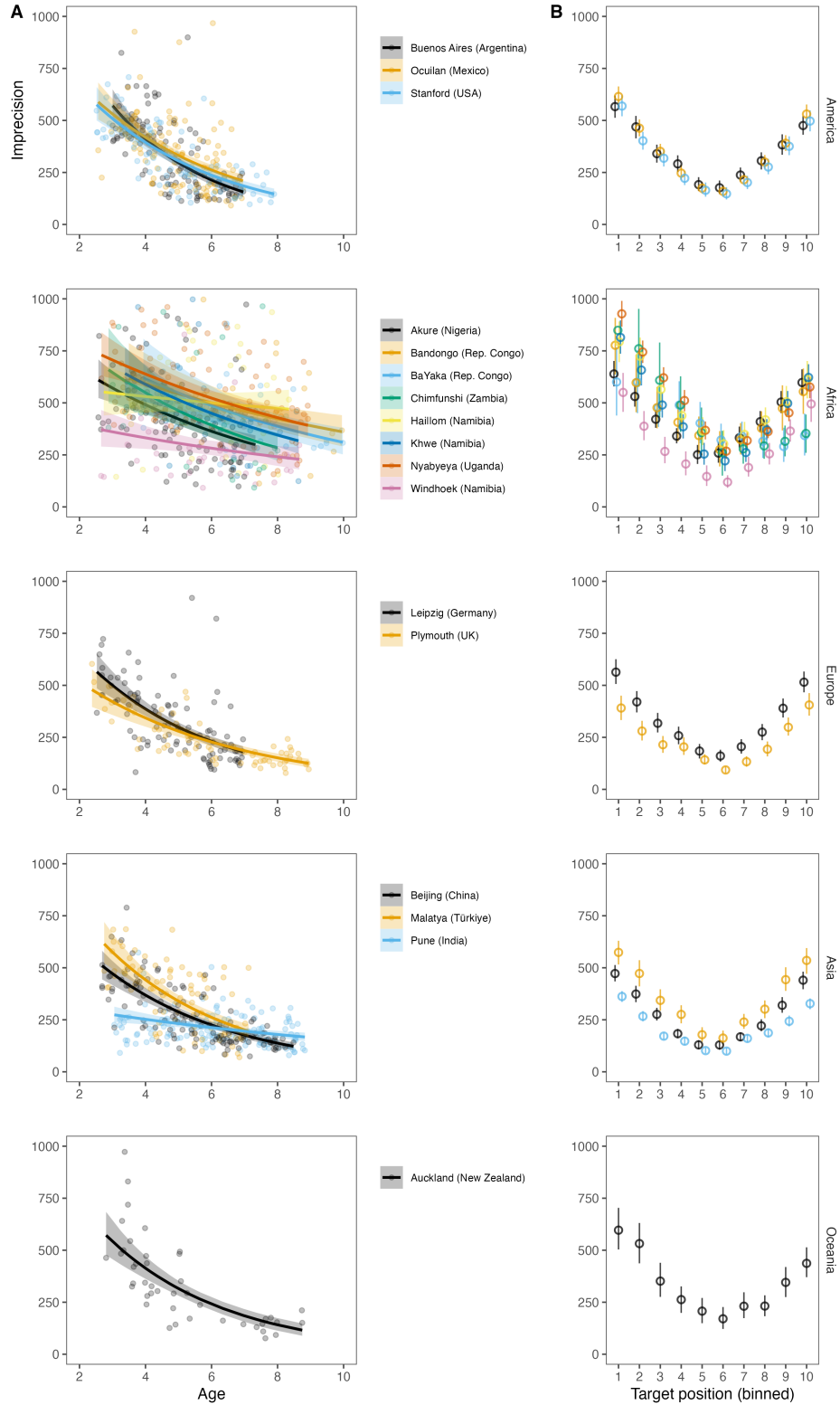
Figure 2. A) Developmental trajectories by cultural settings. Transparent dots show aggregated data for each individual. The color dentoes the different cultural settings. The developmental trajectories are predicted based on a model of the data aggregated for each participant. B) Performance by screen section and cultural setting. Each bin covers 1/10th of the screen. Points show means and eroor bars 95% confidence intervals for the data within

96 is spread over a wider area.

97   The processing signature was clearly visible across all 17 communities. Precision

98 decreased when the agent looked at locations further away from the center (fixed effect: β

99 $= 0.47$, 95% HDI (- 0.54); range of community-level effects: $\beta_{min} = 0.58$, 95% HDI (0.51 -

100 0.66) to $\beta_{max} = 0.16$, 95% HDI (-0.01 - 0.34). Visualization of the data showed the

101 predicted u-shaped pattern in all communities (see Fig. 2B). These results indicate a

102 universal cognitive process used by children in all communities. There is, however, an

103 alternative way in which the u = shaped pattern might arise: it would also arise when

104 participants ignored the agent's gaze and instead always selected the middle of the screen

105 (center bias). To rule out this alternative explanation, we directly compared three cognitive

106 models that made different assumptions about how participants' responses were generated:

107 the focal vector-based gaze estimation model described above, a center-bias model where

108 participants always select the center, and a random guessing model where participants

109 select random locations. For every community, we found overwhelming support for the

110 gaze estimation model (min $BF_{10} > 1000$ for comparisons with both alternative models).

111 Taken together, children from all 17 communities processed gaze in similar ways.

112   What is left to explain are the marked community- and individual-level differences.

113 In addition to the gaze-following task, caregivers filled out a small questionnaire about

114 children's access to screen-based technology and the composition of their households. On

115 an individual level, we found that across communities, children with access to touchscreen

116 devices had higher levels of precision ($\beta =$, SE = , 95% CrI = -). On a community level,

117 we also saw that average performance was lowest in the communities in which touchscreen

118 devices were the least frequent (Fig. 3). Thus, methodological factors likely explained the

119 marked differences between communities.

120   Based on this, we also computed the number of older children living in the household.

121 We did not pre-register this, but decided to also include this variable because prior work on
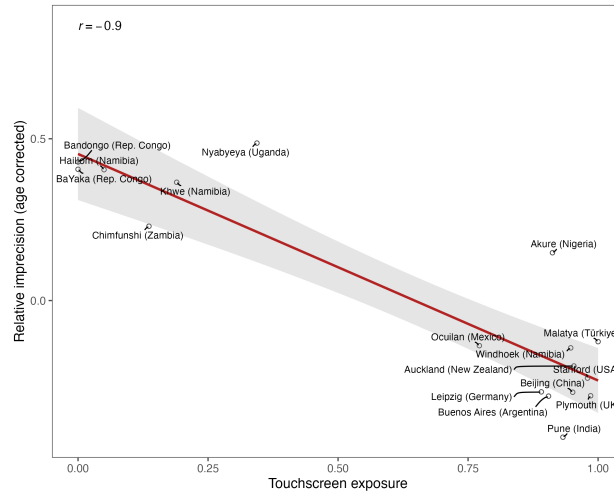
Figure 3. Correlation between performance in the gaze-following task and the exposure to touchscreens on a cultural setting-level. Realtive imprecision accounts for differences in age between settings in that it corresponds to the random intercept estimate for each setting in a model predicting performance by age.

advanced forms of social cognition has suggested that the presence of older siblings is slightly more predictive of development compared to other aspects of household composition (e.g., *21*, *22*, but see *23*). We think that gaze following improves though practice during social interaction. Particularly supportive, we believe, are situations in which it is difficult to infer the goals and intentions of another person based on other information sources (e.g., language or context). In such situations, children have to pay closer attention to the interaction partner and use their gaze to make such inferences. We are well aware that the numbers in the questionnaire are very far removed from these forms of social interaction. Thus, at best, they are a very coarse approximation. As a consequence, we do not expect there to be a very strong link between the two. We sincerely hope that future work will be able to directly assess children's social interactions.

On an individual level, other factors likely generated individual differences because there was substantial variation even in communities where almost all children had access to touchscreens. Previous work suggested that social cognition develops in social interaction

136   (24–26), in particular, with older siblings (21, 22). We found tentative support for this

137   idea in the form of a small but consistent effect in the direction that participants living

138   together with more children older than themselves (relative to other children from the same

139   setting) had slightly higher levels of precision (β = , SE = , 95% CrI = -).

140   Following and understanding gaze is a foundational building block of human social

141   cognition (8, 9). A substantial body of work has explored the developmental onset of gaze

142   following in a few selected cultural settings (10–12, 15, 16, 27). The data reported here

143   provides evidence that children all over the world process gaze in the same way. Key

144   performance signatures of a model that sees gaze following as a form of social vector

145   estimation were found in the data of all 17 distinct cultural communities. The cognitive

146   processes underlying gaze following are thus likely to be rooted in humans' evolved

147   cognitive architecture, which is – presumably – later refined during social interaction

148   (28–30). The phylogenetic roots of these processes might possibly lie much deeper as

149   primates from a wide range of species follow gaze (31–34). Yet, if they also show the same

150   processing signatures has yet to be explored.

151   There are important limitations to this study. The methodological factors that

152   influenced performance might have overshadowed individual and community-level

153   differences that originate from other sources. Importantly, this does not affect our

154   interpretation of the data as evidence for shared cognitive processing because the key

155   processing signatures were present in all communities. The role that social interaction

156   plays as a potential driver of development and source of individual differences is vastly

157   under-explored in the current study. The small effect of the exposure to older children does

158   not indicate which aspects of social interaction enhance gaze processing. Naturalistic data

159   directly recording social interactions across communities would offer crucial information to

160   close this explanatory gap. Recent work in the field of language acquisition has shown how

161   technological innovations can be used to close this explanatory gap (35, 36).

162   The evidence presented here holds far-reaching implications. Our work pioneers a

163 methodological approach that introduces solid individual-level measurement to the

164 cross-cultural study of cognitive development. As such, it serves as a blueprint for future

165 research on a broad spectrum of cognitive abilities. Most importantly, this study offers a

166 much-needed empirical foundation for theories on human nature. The finding that children

167 from diverse cultures deploy similar cognitive processes in interpreting gaze points to an

168 evolutionary basis of basic social cognition, which is refined during development,

169 presumably through social interaction.

References

1. M. Tomasello, M. Carpenter, J. Call, T. Behne, H. Moll, Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences* **28**, 675–691 (2005).

2. K. Laland, A. Seed, Understanding human cognitive uniqueness. *Annual Review of Psychology* **72**, 689–716 (2021).

3. H. M. Wellman, *Making Minds: How Theory of Mind Develops* (Oxford University Press, 2014).

4. J. Henrich, *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter* (princeton University press, 2016).

5. M. Tomasello, H. Rakoczy, What makes human cognition unique? From individual to shared to collective intentionality. *Mind & language* **18**, 121–147 (2003).

6. C. H. Legare, The development of cumulative cultural learning. *Annual Review of Developmental Psychology* **1**, 119–147 (2019).

7. S. V. Shepherd, Following gaze: Gaze-following behavior as a window into social cognition. *Frontiers in integrative neuroscience* **4**, 5 (2010).

8. M. Tomasello, B. Hare, H. Lehmann, J. Call, Reliance on head versus eyes in the gaze following of great apes and human infants: The cooperative eye hypothesis. *Journal of human evolution* **52**, 314–320 (2007).

9. M. Scaife, J. S. Bruner, The capacity for joint visual attention in the infant. *Nature* **253**, 265–266 (1975).

10. Y. Tang, M. R. Gonzalez, G. O. Deák, The slow emergence of gaze-and point-following: A longitudinal study of infants from 4 to 12 months. *Developmental Science*, e13457 (2023).

11. G. Gredebäck, L. Fikke, A. Melinder, The development of joint visual attention: A longitudinal study of gaze following during interactions with mothers and strangers. *Developmental science* **13**, 839–848 (2010).

12. K. Byers-Heinlein, R. K.-Y. Tsui, D. Van Renswoude, A. K. Black, R. Barr, A. Brown, M. Colomer, S. Durrant, A. Gampe, N. Gonzalez-Gomez, others, The development of gaze following in monolingual and bilingual infants: A multi-laboratory study. *Infancy* **26**, 4–38 (2021).

13. R. Brooks, A. N. Meltzoff, The development of gaze following and its relation to language. *Developmental science* **8**, 535–543 (2005).

14. M. Carpenter, K. Nagell, M. Tomasello, G. Butterworth, C. Moore, Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the society for research in child development*, i–174 (1998).

15. M. Hernik, T. Broesch, Infant gaze following depends on communicative signals: An eye-tracking study of 5-to 7-month-olds in vanuatu. *Developmental science* **22**, e12779 (2019).

16. T. Callaghan, H. Moll, H. Rakoczy, F. Warneken, U. Liszkowski, T. Behne, M. Tomasello, W. A. Collins, Early social cognition in three cultural contexts. *Monographs of the society for research in child development*, i–142 (2011).

17. J. C. Prein, S. Kalinke, D. B. Haun, M. Bohn, TANGO: A reliable, open-source, browser-based task to assess individual differences in gaze understanding in 3 to 5-year-old children and adults. *Behavior Research Methods*, 1–17 (2023).

18. P.-C. Bürkner, Bayesian item response modeling in r with brms and stan. *arXiv preprint arXiv:1905.09501* (2019).

19. R. Stengelin, R. Ball, L. Maurits, P. Kanngiesser, D. B. Haun, Children over-imitate adults and peers more than puppets. *Developmental Science* **26**, e13303 (2023).

20. J. C. Prein, L. Maurits, A. Werwach, D. B. M. Haun, M. Bohn, Variation in gaze understanding across the life span: A process-level perspective. *PsyArXiv*, doi: 10.31234/osf.io/dy73a (2023).

21. C. Lewis, N. H. Freeman, C. Kyriakidou, K. Maridaki-Kassotaki, D. M. Berridge, Social influences on false belief access: Specific sibling influences or general apprenticeship? *Child development* **67**, 2930–2947 (1996).

22. T. Ruffman, J. Perner, M. Naito, L. Parkin, W. A. Clements, Older (but not younger) siblings facilitate false belief understanding. *Developmental psychology* **34**, 161 (1998).

23. C. Peterson, V. Slaughter, Opening windows into the mind: Mothers' preferences for mental state explanations and children's theory of mind. *Cognitive Development* **18**, 399–429 (2003).

24. J. Barresi, C. Moore, Intentional relations and social understanding. *Behavioral and brain sciences* **19**, 107–122 (1996).

25. J. Perner, T. Ruffman, S. R. Leekam, Theory of mind is contagious: You catch it from your sibs. *Child Development* **65**, 1228–1238 (1994).

26. H. Rakoczy, Foundations of theory of mind and its development in early childhood. *Nature Reviews Psychology* **1**, 223–235 (2022).

27. C. Moore, The development of gaze following. *Child Development Perspectives* **2**, 66–70 (2008).

28. K. Astor, M. Lindskog, L. Forssman, B. Kenward, M. Fransson, A. Skalkidou, A. Tharner, J. Cassé, G. Gredebäck, Social and emotional contexts predict the development of gaze following in early infancy. *Royal Society open science* **7**, 201178 (2020).

29. J. R. Movellan, J. S. Watson, "The development of gaze following as a bayesian systems identification problem" in *Proceedings 2nd International Conference on Development and Learning. ICDL 2002* (IEEE, 2002), pp. 34–40.

229   30.   A. Senju, A. Vernetti, N. Ganea, K. Hudry, L. Tucker, T. Charman, M. H. Johnson,
            Early social experience affects the development of eye gaze processing. *Current Biol-*
230   *ogy* **25**, 3086–3091 (2015).

231   31.   A. G. Rosati, B. Hare, Looking past the model species: Diversity in gaze-following
            skills across primates. *Current opinion in neurobiology* **19**, 45–51 (2009).
232

233   32.   M. Tomasello, J. Call, B. Hare, Five primate species follow the visual gaze of con-
            specifics. *Animal behaviour* **55**, 1063–1069 (1998).
234

235   33.   S. Itakura, Gaze-following and joint visual attention in nonhuman animals, *Japanese*
            Psychological Research. **46** (2004)pp. 216–226.
236

237   34.   F. Kano, J. Call, Cross-species variation in gaze following and conspecific preference
            among great apes, human infants and adults. *Animal Behaviour* **91**, 137–150 (2014).
238

239   35.   E. Bergelson, M. Soderstrom, I.-C. Schwarz, C. F. Rowland, N. Ramirez-Esparza, L.
            R. Hamrick, E. Marklund, M. Kalashnikova, A. Guez, M. Casillas, others, Everyday
            language input and production in 1,001 children from six continents. *Proceedings of*
            the National Academy of Sciences **120**, e2300671120 (2023).
240

241   36.   S. Donnelly, E. Kidd, The longitudinal relationship between conversational turn-
            taking and vocabulary growth in early language development. *Child Development*
            92, 609–625 (2021).
242