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

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

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

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

## Author Note

Add complete departmental affiliations for each author here. Each new line herein must be indented, like this line.

Enter author note here.

The authors made the following contributions. Manuel Bohn: Conceptualization, Writing - Original Draft Preparation, Writing - Review & Editing; Julia Prein: Conceptualization, Methodology, Writing - Review & Editing; . . . : Conceptualization, Methodology, Writing - Review & Editing; Daniel B. M. Haun: Conceptualization, Writing - Review & Editing.

Correspondence concerning this article should be addressed to Manuel Bohn, Postal 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 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 XXXX participants (age ) in the study lived in 17 different communities spread over five continents. 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 (Fig. 1).

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

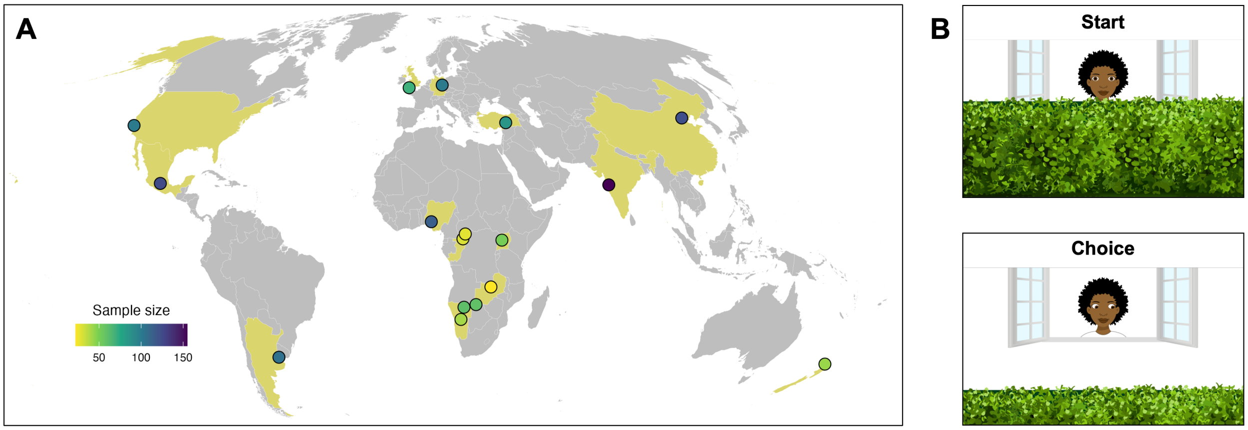


*Figure 1*. (B) Screenshot from the task. The scene depicts the choice phase in a test trial. Participants had to use the gaze of the agent to locate the balloon and click on the hedge where they thought the balloon was. Agents, audio recordings and backgrounds were adpated to each cultural setting.

specifically to each of the 17 communities (Fig. 1). Previous work demonstrated excellent individual-level measurement properties for this task (*17*). Thus, in addition to group-level trends, we were able to investigate individual-level variation.

As the first step, we investigated developmental gains. Across all 17 communities, we found a substantial increase in average levels of precision with age (range: $\beta_{min} =$ , 95% CrI () to $\beta_{max} =$ ()). There were marked differences between communities: highly market-integrated communities around the globe showed higher levels of precision compared to less market-integrated communities. However, we think that these results are better explained by exposure to technology than market integration. We discuss them in more detail below. Differences between communities were small compared to differences between individuals: communities did not form homogeneous clusters but largely overlapping distributions in that some individuals from communities with a lower average level of precision performed better compared to some individuals from a setting with a very high average level of precision. Furthermore, in all communities, some 4-year-olds outperformed children two years older than them (see Fig. 2).

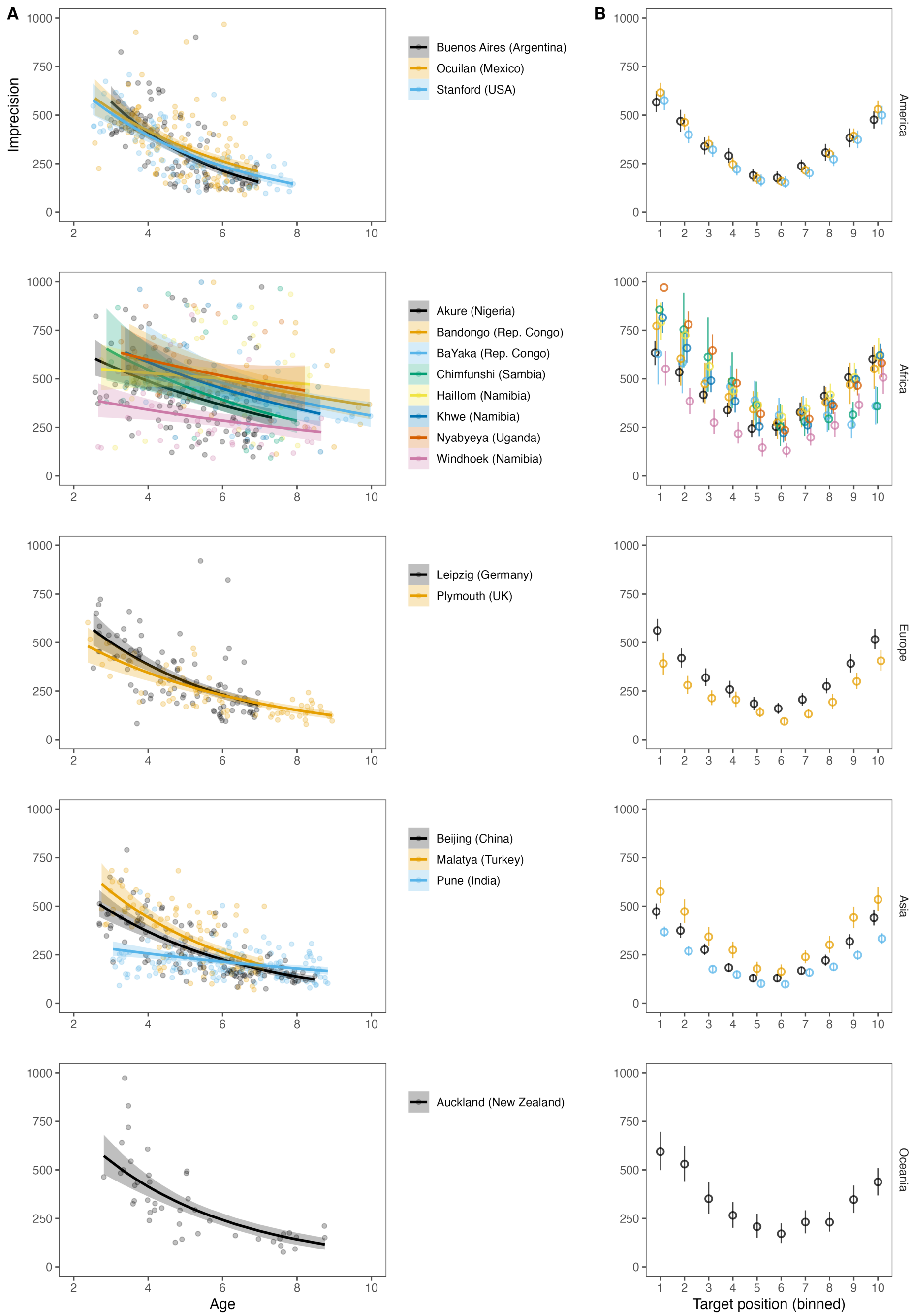


*Figure 2.* A) Developmental trajectories by cultural settings. Transparent dots show aggregated data for each individual. The color dentoes the different cultural settings. The developmental trajectories are predicted based on a model of the data aggregated for each participant. B) Performance by screen 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 that bin aggregated across participants.

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 | 125 (61) | 4.97 (2.57 - 6.95) | Spanish (Mexican) | medium | 0.77 |
| | USA | Stanford | 100 (56) | 5.00 (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 | 41 (18) | 5.68 (2.66 - 8.66) | English (Nigerian) | high | 0.95 |
| | Nigeria | Akure | 114 (53) | 5.07 (2.57 - 7.33) | English (Nigerian) | high | 0.91 |
| | Rep. Congo | BaYaka | 28 (12) | 7.86 (3.94 - 10.56) | BaYaka | low | 0.00 |
| | | Bandongo | 30 (11) | 7.45 (3.50 - 10.95) | Lingala | low | 0.00 |
| | Sambia | Chimfunshi | 22 (5) | 5.98 (2.88 - 8.00) | Bemba | medium | 0.14 |
| | Uganda | Nyabyeya | 51 (27) | 5.41 (3.27 - 8.21) | Swahili | medium | 0.49 |
| 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 | 155 (75) | 6.14 (3.06 - 8.83) | English (Indian) / Marathi | high | 0.92 |
| | Turkey | 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 |

Consistent developmental gains alone cannot inform us about the cognitive processes children use when locating the agent's focus of attention. Recent computational work modeled gaze following as social vector estimation (*18*). When observing the eyes, onlookers estimate a vector from the center of the eye through the pupil. The focus of attention is the location wherever the estimated vectors from both eyes hit a surface (Fig. XX). The estimation process is not perfect, but each individual has a systematic level of uncertainty, which is conceptualized as the cause of individual differences. Importantly, this process model predicts a clear performance signature in our cross-cultural gaze following task: Trials in which the agent looks further away from the center (i.e. to the left or right side of the screen) should result in lower levels of precision, compared to trials in which the agent looks closer to the middle. This prediction is best understood by considering a similar phenomenon: pointing a torch light to a flat surface. The width of the light beam represents each individual's level of uncertainty in vector estimation. When the torch is directed straight down, the light beam is concentrated in a relatively small area. When the torch is rotated to the side, the light from one half of the cone must travel further than the light from the other half to reach the surface. As a consequence, the light is spread over a wider area.

The processing signature was clearly visible across all 17 communities. Precision decreased when the agent looked at locations further away from the center (fixed effect: $\beta =$ (), community-level effects: Fig. 2). Visualization of the data showed the predicted u-shaped pattern in all communities. These results indicate a universal cognitive process used by children in all communities. There is, however, an alternative way in which the u = shaped pattern might arise: it would also arise when participants ignored the agent's gaze and instead always selected the middle of the screen (center bias). To rule out this alternative explanation, we directly compared three cognitive models that made different assumptions about how participants' responses were generated: the focal vector-based gaze estimation model described above, a center-bias model where participants always select the center, and a random guessing model where participants select random locations. For every community, we found overwhelming support for the gaze estimation model (min $BF_{10} < 1000$ for comparisons with both alternative models). Taken together, children from all 17 communities processed gaze in similar ways.

What is left to explain are the marked community- and individual-level differences. In addition to the gaze-following task, caregivers filled out a small questionnaire about children's access to screen-based technology and the composition of their households. On an individual level, we found that across communities, children with access to touchscreen devices had higher levels of precision (fixed effect: $\beta =$ (), community-level effects: Fig. 2). On a community level, we also saw that average performance was lowest in the communities in which touchscreen devices were the least frequent (Fig. 3). Thus, methodological factors likely explained the marked differences between communities.

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 (*19–21*), in particular, with older siblings (*22*, *23*). We found tentative support for this idea in the form of a small but consistent effect in the direction that participants living together with more children older than themselves had slightly higher levels of precision (fixed effect: $\beta =$ (), community-level effects: Fig. 2).

Following and understanding gaze is a foundational building block of human social cognition (*8*, *9*). A substantial body of work has explored the developmental onset of gaze following in a few selected cultural settings (*10–12*, *15*, *16*). The data reported here provides evidence that children all over the world process gaze in the same way. Key performance signatures of a model that sees gaze following as a form of social vector estimation were found in the data of all 17 distinct cultural communities. The cognitive processes underlying gaze following are thus likely to be rooted in humans' evolved cognitive architecture, which is – presumably – later refined during social interaction. The phylogenetic roots of these processes might possibly lie much deeper as primates from a wide range of species follow gaze (*24–26*). Yet, if they also show the same processing signatures has yet to be explored.

There are important limitations to this study. The methodological factors that influenced performance might have overshadowed individual and community-level differences that originate from other sources. Importantly, this does not affect our interpretation of the data as evidence for shared cognitive processing because the key processing signatures were present in all communities. The role that social interaction plays as a potential driver of development and source of individual differences is vastly under-explored in the current study. The small effect of the exposure to older children does not indicate which aspects of social interaction enhance gaze processing. Naturalistic data directly recording social interactions across communities would offer crucial information to close this explanatory gap.

Recent work in the field of language acquisition has shown how technological innovations can be used to close this explanatory gap.

The evidence presented here holds far-reaching implications. Our work pioneers a methodological approach that introduces solid individual-level measurement to the cross-cultural study of cognitive development. As such, it serves as a blueprint for future research on a broad spectrum of cognitive abilities. Most importantly, this study offers a much-needed empirical foundation for theories on human nature. The finding that children from diverse cultures deploy similar cognitive processes in interpreting gaze points to an evolutionary basis of basic social cognition, which is refined during development, 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. 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).

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

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

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

22. 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).

23. 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).

24. 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).

25. M. Tomasello, J. Call, B. Hare, Five primate species follow the visual gaze of conspecifics. *Animal behaviour* **55**, 1063–1069 (1998).

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