A cross-cultural universal of human social cognition: Children from 17 diverse communities process gaze in similar ways

Manuel Bohn (ORCID: 0000-0001-6006-1348)[1,2,*], Julia Prein (ORCID: 0000-0002-3154-6167)[1,2,*], Agnes Ayikoru[3], Florian M. Bednarski (ORCID: 0000-0003-4384-4791)[4], Ardain Dzabatou[5], Michael C. Frank (ORCID: 0000-0002-7551-4378)[6], Annette M. E. Henderson (ORCID: 0000-0003-4384-4791)[4], Joan Isabella[3], Josefine Kalbitz[2], Patricia Kanngiesser (ORCID:0000-0003-1068-3725)[7], Dilara Keşşafoğlu (ORCID: 0000-0002-7356-0733)[8], Bahar Köymen (ORCID: 0000-0001-5126-8240)[9], Maira V. Manrique-Hernandez[2], Shirley Magazi (ORCID: 0009-0006-0479-9800)[10], Lizbeth Mújica-Manrique[2], Julia Ohlendorf[2], Damilola Olaoba[2], Wesley R. Pieters (ORCID:0000-0002-6152-249X)[10], Sarah Pope-Caldwell[2], Katie Slocombe (ORCID: 0000-0002-7310-1887)[11], Robert Z. Sparks (ORCID: 0000-0001-7545-0522)[6], Jahnavi Sunderarajan[2], Wilson Vieira[2], Zhen Zhang (ORCID: 0000-0001-9300-0920)[12], Yufei Zong (ORCID: 0009-0000-5012-0244)[12], Roman Stengelin (ORCID: 0000-0003-2212-4613)[2,10], & 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

[3] Budongo Conservation Field Station

[4] School of Psychology, University of Auckland

[5] Université Marien Ngouabi
[6] Department of Psychology, Stanford University
[7] School of Psychology, University of Plymouth
[8] Department of Psychology, Koç University
[9] Division of Psychology, Communication, and Human Neuroscience, University of Manchester
[10] Department of Psychology and Social Work, University of Namibia
[11] Department of Psychology, University of York
[12] CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences
* joint first author

Author Note

The authors would like to thank Luke Maurits for statistical advice. Manuel Bohn was supported by a Jacobs Foundation Research Fellowship (2022-1484-00). We are greatful to thank all children and caregivers for participating in the study.

The authors made the following contributions. Manuel Bohn (ORCID: 0000-0001-6006-1348): Conceptualization, Methodology, Formal Analysis, Writing - Original Draft Preparation, Writing - Review & Editing; Julia Prein (ORCID: 0000-0002-3154-6167): Conceptualization, Methodology, Software, Investigation, Writing - Review & Editing; Agnes Ayikoru: Investigation; Florian M. Bednarski (ORCID: 0000-0003-4384-4791): Investigation, Writing - Review & Editing; Ardain Dzabatou: Investigation; Michael C. Frank (ORCID: 0000-0002-7551-4378): Investigation, Writing - Review & Editing; Annette M. E. Henderson (ORCID: 0000-0003-4384-4791): Investigation, Writing - Review & Editing; Joan Isabella: Investigation; Josefine Kalbitz: Investigation, Writing - Review & Editing; Patricia Kanngiesser (ORCID:0000-0003-1068-3725): Investigation, Writing - Review & Editing; Dilara Keşşafoğlu (ORCID: 0000-0002-7356-0733): Investigation, Writing - Review & Editing; Bahar Köymen (ORCID: 0000-0001-5126-8240): Investigation, Writing - Review & Editing; Maira V. Manrique-Hernandez: Investigation; Shirley Magazi (ORCID: 0009-0006-0479-9800): Investigation; Lizbeth Mújica-Manrique: Investigation, Writing - Review & Editing; Julia Ohlendorf: Investigation; Damilola Olaoba: Investigation; Wesley R. Pieters (ORCID:0000-0002-6152-249X): Investigation, Writing - Review & Editing; Sarah Pope-Caldwell: Investigation; Katie Slocombe (ORCID: 0000-0002-7310-1887): Investigation; Robert Z. Sparks (ORCID: 0000-0001-7545-0522): Investigation; Jahnavi Sunderarajan: Investigation; Wilson Vieira: Investigation; Zhen Zhang (ORCID: 0000-0001-9300-0920): Investigation, Writing - Review & Editing; Yufei Zong (ORCID: 0009-0000-5012-0244): Investigation, Writing - Review & Editing; Roman Stengelin (ORCID: 0000-0003-2212-4613): Conceptualization, Methodology, Investigation, Writing -

Review & Editing; Daniel B. M. Haun: Conceptualization, Funding acquisition, Writing - Review & Editing.

Correspondence concerning this article should be addressed to Manuel Bohn (ORCID: 0000-0001-6006-1348), Universitätsallee 1, 21335 Lüneburg, Germany. E-mail: manuel.bohn@leuphana.de

Abstract

Theoretical accounts assume that key features of human social cognition are universal. Here we focus on gaze following, the bedrock of social interactions and coordinated activities, to test this claim. In this comprehensive cross-cultural study spanning five continents and 17 distinct cultural communities, we examined the development of gaze following in early childhood. We identified key processing signatures through a computational model that assumes that participants follow an individual's gaze by estimating a vector emanating from the eye-center through the pupil. Using a single reliable touchscreen-based task, we found these signatures in all communities, suggesting that children worldwide processed gaze in highly similar ways. Additionally, we found a cross-culturally consistent relationship between children's exposure to touchscreens and their performance in the task, which also explained absolute differences between communities. These results provide strong evidence for a universal process underlying a foundational socio-cognitive ability in humans that can be reliably inferred even in the presence of cultural variation in overt behavior.

Keywords: keywords

Word count: X

A cross-cultural universal of human social cognition: Children from 17 diverse communities process gaze in similar ways

Human socio-cognitive skills enable unique forms of communication and cooperation that provide a bedrock for cumulative culture and the formation of complex societies (*1–7*). The eyes are the proverbial "window to the mind" and eye gaze is essential for many social reasoning processes (*8*, *9*). Others' eye gaze is used to infer their focus of visual attention, which is a critical aspect of coordinated activities, including communication and cooperation (*10*, *11*). The ability to follow gaze emerges early in childhood (*12–15*) and individual differences in children's gaze-following ability predict later life outcomes, most notably, communicative abilities (*16*, *17*). Difficulties with gaze following have been linked to developmental disorders, including Autism (*18*).

There is a widely-held assumption, that, despite substantial variation in developmental contexts, gaze-following works and develops in the same way across human societies (*19*). However, this view is seemingly at odds with the substantial diversity in socio-cognitive development that cross-cultural studies have revealed (*3*, *20–23*). One potential source for this paradox lies in the reliance on aggregated measures in cross-cultural studies. Absolute differences in mean performance across communities are interpreted as a signal of different underlying cognitive processes. In this study, we resolve this paradox by instead focusing on processing signatures that can be investigated independently of absolute community-level differences. This allows us to directly evaluate the empirical foundation of claims about universal features of human social cognition. To this end, we conducted a pre-registered, large-scale, cross-cultural study on the development of gaze-following abilities to study potentially universal processing signatures. These signatures were derived from a computational model that assumes that participants follow gaze by estimating a vector emanating from the eye center through the pupil (*24*).

The 1377 participants who took part in the study lived in 17 different communities across 14 countries and five continents (Fig. 1A, Tab. 1). These countries represent ~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). This diversity allowed us to overcome the common pitfall of cross-cultural studies that largely compared urban communities from the global north to rural communities from the global south (*25*).

Table 1

Participant demographics.

Continent	Country	Community	N(male)	Age (range)	Language	Market integration	Touchscreen
Americas	Argentina	Buenos Aires	105 (53)	4.72 (3.00 - 6.96)	Spanish (Rioplatense)	high	0.90
	México	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

Table 1 continued

Continent	Country	Community	N(male)	Age (range)	Language	Market integration	Touchscreen
		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)	Swahili	medium	0.34
	Zambia	Chimfunshi	22 (5)	5.98 (2.88 - 8.00)	Bemba	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

Table 1 continued

Continent	Country	Community	N(male)	Age (range)	Language	Market integration	Touchscreen
	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

Note. *Local collaborators and piloting suggested that Nigerian English is suitable for Windhoek as well.

We used an animated picture book tablet task in which participants had 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 (im)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 the child's response. We adapted visuals and audio instructions specifically to each of the 17 communities. Previous work demonstrated excellent individual-level measurement properties for this task in a German sample (*26*).

As the first step, we investigated developmental improvements, that is, how children become more precise at estimating the target location with age. Across all 17 communities,

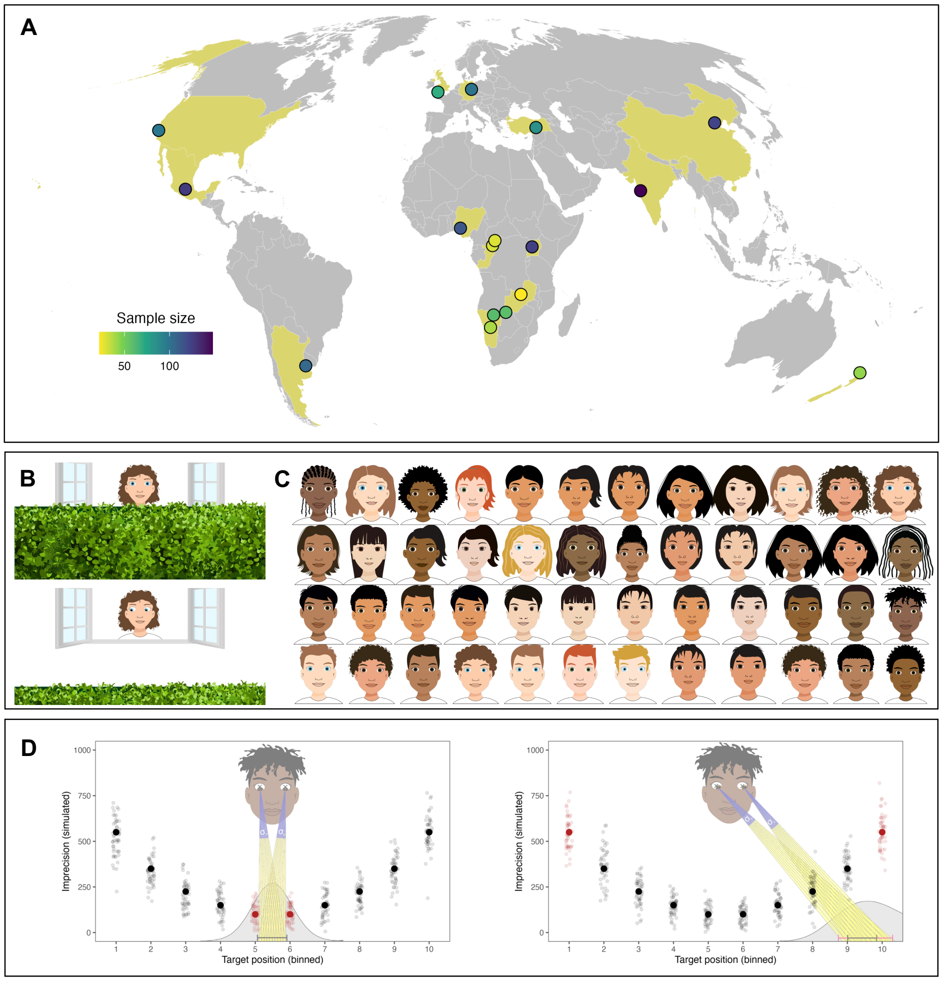

Figure 1. (A) Data collection sites. Points show the approximate geographical location of the data collection sites, coloring shows the sample sizes. (B) Screenshot from the task. The scene depicts the choice phase in a test trial. Participants had to use the gaze of the agent to locate the balloon and click on the location on the hedge where they thought the balloon was. Agents, audio recordings and backgrounds were adapted 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 (transparent lines). Individuals differ in their level of precision (indicated by sigma). For a given level of precision, the closer the target lands to the left or the right, the less precise the model predicts individuals to be. Solid and transparent dots show simulated means and individual data points to illustrate the predicted effect of target position.

we found a substantial increase in average levels of precision with age (fixed effect in Bayesian regression model (*27*): β = -0.30, 95% HDI (-0.40 - -0.21); range of community-level (random) effects: β_{min} = -0.06, 95% HDI (-0.18 - 0.05) to β_{max} = -0.59, 95% HDI (-0.71 - -0.48).

Nevertheless, there were also marked differences between communities (see Fig. 2A). In a six-fold cross-validation procedure, we trained a regression model on a subset of the data (training data) to later predict the held-out data (testing data) (*28*). We found that a model assuming cross-cultural variation in average performance as well as cross-cultural variation in developmental trajectories outperformed simpler models – assuming no variation in the shape of developmental trajectories or nor variation between settings at all – in 98% of cases (see Supplemental Material). At first glance, it seems that highly market-integrated communities around the globe showed higher levels of precision compared to less market-integrated communities (compare Tab. 1 and Fig. 2A). However, numerous alternative groupings (e.g., average levels of education or average household size) are possible and, instead, these results could be best understood in terms of exposure to touch-screen devices; a finding we discuss in more detail below. Importantly, average differences in precision 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. Similarly, in all communities, some 4-year-olds outperformed children two years older than them (see Fig. 2A). The lack of adequate individual-level measurement instruments in previous large-scale developmental cross-cultural studies made it impossible to contrast these perspectives. The substantial overlap between communities found here speaks against categorical differences in gaze following and is suggestive of a similar underlying process. However, consistent developmental improvements alone cannot inform us about the cognitive processes children use when locating the agent's focus of attention.

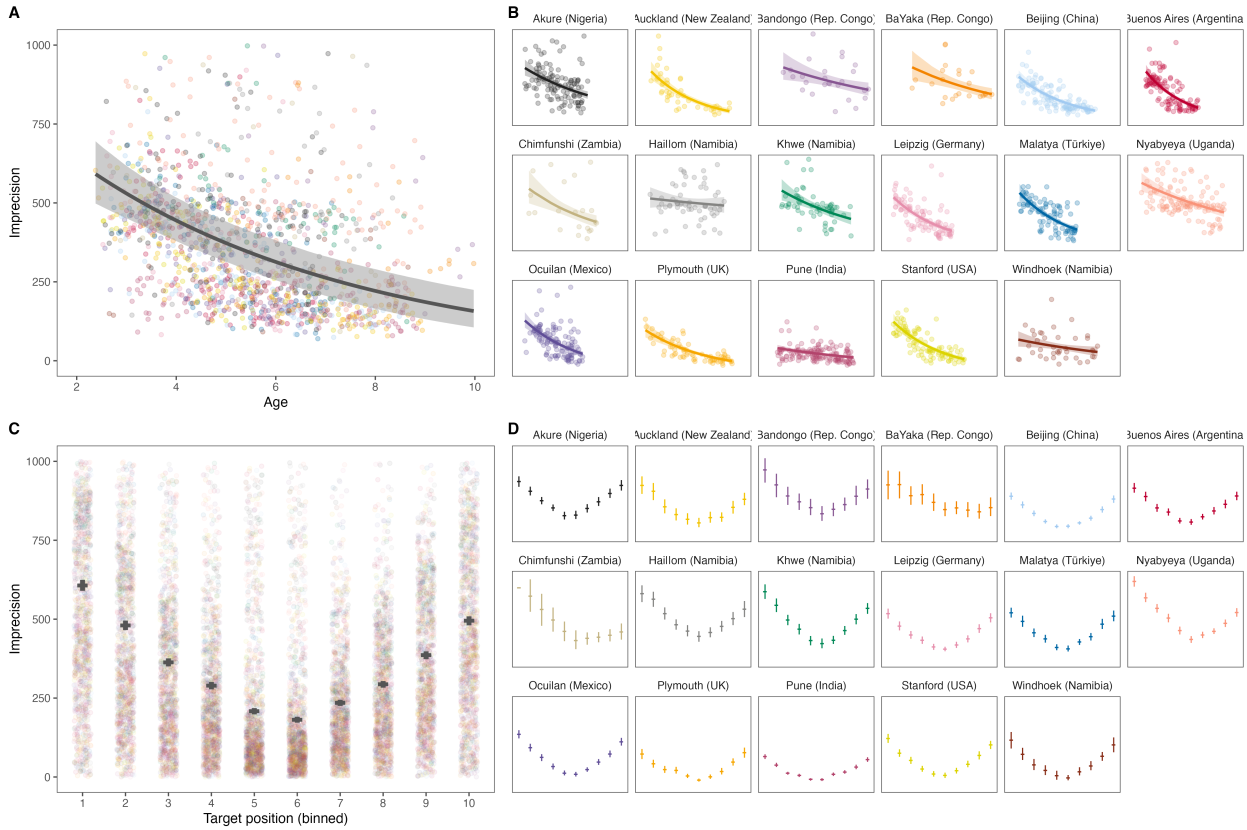

Figure 2. A) Developmental trajectory across and B) by community. The developmental trajectories are predicted based on a model of the data aggregated for each participant. C) Performance by screen section across and D) by community. Each bin covers 1/10th of the screen. Points show means, and error bars 95% confidence intervals for the data within that bin aggregated across participants. Transparent dots in A) and C) show aggregated data for each individual.

Recent computational work modeled gaze following as social vector estimation (*24*). When observing the eyes, onlookers estimate a vector emanating from the center of the eye that runs through the pupil. The focus of attention is the location where the estimated vectors from both eyes hit a surface (Fig. 1D). It is assumed that this estimation process has some uncertainty and that individuals vary in their level of uncertainty. As a consequence, even though individuals use the same general process, they might differ in their absolute levels of precision. Crucially, this process model predicts a clear performance signature in our 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 (see Fig. 1D).

This processing signature was clearly visible across all 17 communities. Precision decreased when the agent looked at locations further away from the center (fixed effect: β = 0.47, 95% HDI (- 0.54); range of community-level effects: β_{min} = 0.58, 95% HDI (0.51 - 0.66) to β_{max} = 0.16, 95% HDI (-0.01 - 0.33). Visualization of the data showed the predicted u-shaped pattern in all communities (see Fig. 2B). These results indicate a universal cognitive process used by children in all communities. There are, however, alternative ways in which the u-shaped pattern might arise: if participants ignored the agent's gaze and instead always selected the middle of the screen (center bias) or randomly select locations (random guessing), precision would also decrease when the balloon lands further away from the center. To rule out these alternative explanations, we 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 $BF_{10} > 100\ 000$ for comparisons with both alternative models). Taken together, children from all 17 communities processed gaze in similar ways.

Next, we looked at factors that could explain community- and individual-level variation. In addition to the gaze-following task, caregivers responded to a short questionnaire about children's access to screen-based technology and household composition. On an individual level, we found that children with access to touchscreen devices had higher levels of precision (β = -0.14, SE = 0.04, 95% CrI = -0.21 - -0.07). This effect was consistent across communities in that allowing the effect of access to touchscreens to vary across communities did not improve model fit (see Supplemental Materials). On a community level, we also saw that average performance was lowest in communities in which touchscreen devices were the least frequent (community-level correlation between age-corrected imprecision and proportion of children with access to touchscreens: r = -0.90, 95% CI = -0.96 - -0.74). Thus, familiarity with the device used for data collection likely explains variation between communities. The most likely explanation for this effect is that children with more touchscreen experience were better at task handling and thus more likely to precisely touch the location they inferred the agent to look at.

However, there was substantial variation between individuals that could not be explained by differential exposures to touch screens alone. For example, in Malatya (Türkyie) where 100% of children had access to touch screens there was still substantial variation between individuals (see Fig. 1B). This strongly indicates that other factors likely contributed to individual differences. Social interaction has been highlighted as an important driver of social-cognitive development (e.g., *19*, *29–32*) and thus we hypothesized (and pre-registered) that more opportunities for social interaction – approximated by living in larger households with more children – would be associated with higher levels of precision. At first glance, the opposite seemed to be the case: there was a substantial zero-order correlation between the number of children living in the household and average imprecision (0.27, 95% CI = 0.22 - 0.32) suggesting that having more opportunities for social interactions was related to poorer performance. However, this correlation is spurious and reflects a confounding of household composition with exposure to technology on a community level. When predicting performance by relative opportunities for social interactions within a community – while accounting for absolute differences and the prevalence of touchscreens – we found no strong associations between any of the demographic indicators and performance (see Supplemental Material). The reason for an absence of such an association may be a lack of resolution: household composition is very far removed from the factors that previous work has suggested to be related to the development of gaze following in younger children, such as attachment quality or the use of gaze in early communicative interactions (*33–35*). Future work could increase the resolution with which everyday experiences in children from diverse communities are recorded to compare the drivers behind development as we observe it. Recent work in the field of language acquisition has shown how technological innovations allowed for direct recording of social interactions across communities which can be used to close this explanatory gap (*36*, *37*).

Following and understanding gaze is a foundational building block of human social cognition (*10*, *11*). A substantial body of work has explored the developmental onset of gaze following in a few selected cultural communities (*12–14*, *38*). The data reported here provides strong evidence that children from a large and diverse set of communities process others' gaze in similar ways. We found key performance signatures of a model treating gaze following as a form of social vector estimation across all 17 communities. With the focus on individual-level processing signatures, the study goes beyond previous studies on gaze following – focused on the onset of gaze following in infancy (*39*, *40*) – as well as comprehensive cross-cultural studies that compared average developmental trajectories (*41–44*).

The cognitive processes underlying gaze following might be rooted in humans' evolved cognitive architecture, which is – presumably – later refined during social (i.e., cultural) interaction (*33–35*). The phylogenetic roots of these processes might possibly lie much deeper as primates from a wide range of species follow gaze (*45–48*). Yet, similarities in overt behavior do not imply the same underlying cognitive processes. The present study defines clear performance signatures that can be explored in other species to test such evolutionary hypotheses.

Our study combined precise individual-level cognitive measurement and individual-level assessment of experience (here: touchscreen exposure) in a large and diverse sample to directly investigate the impact of specific cultural experiences on developmental outcomes. Instead of establishing universality by maximizing the cultural distance between two or three tested communities (*49*), this large-scale cross-cultural approach treats children's cultural experience at scale, shedding light on the big "middle ground" of children's cultural experience (*25*).

In sum, our work pioneers an approach that introduces computational modeling and precise individual-level measurement to the cross-cultural study of cognitive development. This approach allowed us to test for universals in the human cognitive architecture rather than just overt behavior. As such, it can serve as a blueprint for future research on a broad spectrum of cognitive abilities and offers a much-needed empirical foundation for theories on the nature of the human mind. Children from diverse cultures deploy similar cognitive processes in interpreting gaze, pointing to a universal foundation of basic social cognition, which is refined during development.

References

1. M. Tomasello, The adaptive origins of uniquely human sociality. *Philosophical Transactions of the Royal Society B* **375**, 20190493 (2020).

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. C. Heyes, *Cognitive Gadgets* (Harvard University Press, 2018).

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

9. M. J. Doherty, The development of mentalistic gaze understanding. *Infant and Child Development* **15**, 179–186 (2006).

10. 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).

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

12. 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).

13. 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).

14. 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).

15. T. Del Bianco, T. Falck-Ytter, E. Thorup, G. Gredebäck, The developmental origins of gaze-following in human infants. *Infancy* **24**, 433–454 (2019).

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

17. 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).

18. R. J. Itier, M. Batty, Neural bases of eye and gaze processing: The core of social cognition. *Neuroscience & Biobehavioral Reviews* **33**, 843–863 (2009).

19. M. Tomasello, *Becoming Human: A Theory of Ontogeny* (Harvard University Press, 2019).

20. J. G. Miller, M. Wice, N. Goyal, Contributions and challenges of cultural research on the development of social cognition. *Developmental Review* **50**, 65–76 (2018).

21. A. Mayer, B. E. Träuble, Synchrony in the onset of mental state understanding across cultures? A study among children in samoa. *International Journal of Behavioral Development* **37**, 21–28 (2013).

22. H. G. Dixson, A. F. Komugabe-Dixson, B. J. Dixson, J. Low, Scaling theory of mind in a small-scale society: A case study from vanuatu. *Child Development* **89**, 2157–2175 (2018).

23. M. Taumoepeau, S. Sadeghi, A. Nobilo, Cross-cultural differences in children's theory of mind in iran and new zealand: The role of caregiver mental state talk. *Cognitive Development* **51**, 32–45 (2019).

24. 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).

25. H. C. Barrett, Towards a cognitive science of the human: Cross-cultural approaches and their urgency. *Trends in Cognitive Sciences* **24**, 620–638 (2020).

26. 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).

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

28. 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).

29. J. Carpendale, C. Lewis, *What Makes Us Human: How Minds Develop Through Social Interactions* (Routledge, 2020).

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

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

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

33. 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).

34. 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.

35. 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 Biology* **25**, 3086–3091 (2015).

36. 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).

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

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

39. 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).

40. 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).

41. B. R. House, P. Kanngiesser, H. C. Barrett, T. Broesch, S. Cebioglu, A. N. Crittenden, A. Erut, S. Lew-Levy, C. Sebastian-Enesco, A. M. Smith, others, Universal norm psychology leads to societal diversity in prosocial behaviour and development. *Nature Human Behaviour* **4**, 36–44 (2020).

42. E. J. Van Leeuwen, E. Cohen, E. Collier-Baker, C. J. Rapold, M. Schäfer, S. Schütte, D. B. Haun, The development of human social learning across seven societies. *Nature Communications* **9**, 2076 (2018).

43. P. R. Blake, K. McAuliffe, J. Corbit, T. C. Callaghan, O. Barry, A. Bowie, L. Kleutsch, K. Kramer, E. Ross, H. Vongsachang, others, The ontogeny of fairness in seven societies. *Nature* **528**, 258–261 (2015).

44. P. Kanngiesser, M. Schäfer, E. Herrmann, H. Zeidler, D. Haun, M. Tomasello, Children across societies enforce conventional norms but in culturally variable ways. *Proceedings of the National Academy of Sciences* **119**, e2112521118 (2022).

45. 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).

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

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

48. 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).

49. A. Norenzayan, S. J. Heine, Psychological universals: What are they and how can we know? *Psychological Bulletin* **131**, 763 (2005).