



OPEN ACCESS

EDITED BY

Rachel L. Severson,
University of Montana, United States

REVIEWED BY

Charles Beekman,
Independent Researcher, West Springfield, VA,
United States
Shailee Woodard,
Keene State College, United States

*CORRESPONDENCE

Ishanti Gangopadhyay
✉ ishgang@iu.edu

RECEIVED 15 October 2024

ACCEPTED 03 February 2025

PUBLISHED 19 February 2025

CITATION

Gangopadhyay I and Peters L (2025) The effect of speaker reliability on word learning in children: a replication study. *Front. Dev. Psychol.* 3:1469550. doi: 10.3389/fdyps.2025.1469550

COPYRIGHT

© 2025 Gangopadhyay and Peters. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The effect of speaker reliability on word learning in children: a replication study

Ishanti Gangopadhyay* and Lillian Peters

Department of Speech, Language and Hearing Sciences, Indiana University, Bloomington, IN, United States

This study aimed to replicate a previously conducted in-person speaker reliability experiment using a fully online methodology. Twenty children aged 4 to 6 years participated in a live video call with the experimenter and completed virtual tasks on a web-based platform. The experimental task mirrored that used in the previous reliability study, where children learned novel words from both a reliable and an unreliable speaker, followed by testing children's novel word retention. Consistent with the prior study's findings, children performed above chance in both conditions and retained novel labels taught by both speakers. These preliminary results suggest that speaker reliability is a robust cue, showing consistent effects across different data collection methods. Thus, online data collection shows promise for producing viable results and improving participation by making research more accessible and flexible. Nonetheless, further studies are necessary to explore its strengths and limitations, especially in the context of research involving children.

KEYWORDS

speaker reliability, online testing, word learning, children, replication

1 Introduction

Children are sensitive to the traits and behaviors of people around them, and children use these characteristics to evaluate speakers as potential informants (Liu et al., 2013). Researchers have found that children prefer to learn information from speakers that are more familiar to them (Wood et al., 2013), more attractive to them (Bascandziev and Harris, 2016), and have familiar accents (Kinzler et al., 2011), among other attributes. One speaker cue that has received considerable attention over the decades is the speaker's reliability status (Corriveau and Harris, 2009; Koenig et al., 2004; Koenig and Harris, 2005; Nurmsoo and Robinson, 2009; Sobel and Macris, 2013; Sobel et al., 2012). Speaker reliability describes an informant's past accuracy in labeling objects, where a reliable speaker correctly labels objects known to the child, and an unreliable speaker incorrectly labels the same known objects. From a robust literature of word-learning experiments, we know that children prefer to learn names of objects from accurate, reliable speakers rather than from inaccurate, unreliable speakers (Corriveau et al., 2009; Fusaro and Harris, 2008; Jaswal and Neely, 2006; Koenig et al., 2004; Koenig and Woodward, 2010). However, other studies have shown that, despite this speaker preference, children end up retaining novel information from both types of speakers (Gangopadhyay and Kaushanskaya, 2022; Krogh-Jespersen and Echols, 2012; Mangardich and Sabbagh, 2018; Sabbagh and Shafman, 2009).

Traditionally, studies on speaker reliability have utilized various offline, in-person methodologies. However, with the increasing need for alternative data collection methods, such as online testing, especially in light of the COVID-19 pandemic, child language

research must adapt (Hoffmann et al., 2024; Lefever et al., 2007). This change is essential for ensuring that results obtained from online methodologies are as valid and generalizable as those obtained from traditional in-person methods. This effort also aligns with broader concerns of research findings failing to replicate when tested under new conditions (Ioannidis, 2005). Our study focuses on determining the replicability of an in-person speaker reliability task designed by Gangopadhyay and Kaushanskaya (2022; henceforth G&K) using a web-based online platform with preschool children.

1.1 Speaker reliability studies

Studies examining the effects of speaker reliability on word learning in children have revealed insightful patterns regarding how children acquire new words from reliable and unreliable speakers. For instance, research has shown that preschool children prefer to seek out new information from reliable rather than unreliable speakers (Corriveau and Harris, 2009; Cossette et al., 2020; Dautriche et al., 2021; Fusaro and Harris, 2008; Pasquini et al., 2007). Additionally, studies have shown that children not only prefer to seek out information from reliable vs. unreliable speakers but also tend to select names provided by the reliable speaker vs. the unreliable speaker (Fusaro and Harris, 2008; Jaswal and Neely, 2006; Koenig et al., 2004; Koenig and Woodward, 2010). Yet, other studies have demonstrated that children are able to retain novel words from both reliable and unreliable speakers (G&K; Krogh-Jespersen and Echols, 2012; Mangardich and Sabbagh, 2018; Sabbagh and Shafman, 2009).

Several factors may explain these patterns of findings. One is the specific aspect of word learning that was assessed. Some experiments assessed word learning by having children identify a new word-form and select the object from multiple possible referents (e.g., Corriveau et al., 2009), while others assessed word learning by having children maintain and recall that word-object association over time (e.g., G&K; Koenig et al., 2004). Thus, it is possible that reliability studies have captured word learning on different developmental timescales. Second, and relevant to the current study, are differences in the study designs and methodologies employed. Apart from obvious differences in stimuli used among studies, experiments have also differed in how the speaker's reliability was determined (e.g., "blicket" vs. "I think this is a Blicket"), how children were tested on the speaker's knowledge (e.g., "Who gave you all the right names?" vs. "who would you like to ask?," and how children's learning was tested (e.g., match/mismatch vs. retrieval). Additionally, some investigations have used live, in-person speakers to examine how children learned from different speakers (Birch et al., 2008; Jaswal and Neely, 2006; Koenig et al., 2004; Koenig and Woodward, 2010; Krogh-Jespersen and Echols, 2012; Pasquini et al., 2007; Scofield and Behrend, 2008), while some have used computerized tasks where children viewed videos of speakers labeling objects (Corriveau and Harris, 2009; G&K; Koenig and Harris, 2005; Mangardich and Sabbagh, 2018; Sabbagh and Shafman, 2009). Altogether, these findings underscore the methodological diversity employed in investigating speaker reliability and word learning in children. However, all of these

studies were conducted in-person in laboratory settings, and the next crucial step in advancing our knowledge of selective word learning in children is to conduct these types of studies fully online.

1.2 Need for online methods

The landscape of child language research is rapidly evolving, driven by the necessity to adapt to unforeseen challenges and leverage technological advancements (Feijoo et al., 2023; Tomasello, 2020). The recent pandemic highlighted a critical need for alternative data collection methods as in-person studies became impractical or impossible (Hoffmann et al., 2024). This shift toward online methodologies is not just a matter of convenience but a fundamental evolution of how we conduct research in the social sciences, including child language development.

In traditional settings, children's exposure to experimental stimuli and their responses are meticulously controlled, yielding highly reliable data (Fernald et al., 2006). However, these methods are limited by geographical, temporal, and resource constraints. Online testing offers a solution that transcends these barriers, providing access to a broader and more diverse participant pool (Rhodes et al., 2020). This inclusivity is crucial for ensuring that findings are representative of the general population, including underserved and hard-to-reach communities. Moreover, online methodologies enable the collection of data at a scale and speed unattainable in lab settings (Lefever et al., 2007). With sophisticated digital tools, researchers can track engagement, response times, and other metrics with high precision, ensuring that the quality of data is not compromised. These methods also facilitate longitudinal studies, where participants can be easily followed over extended periods, providing richer insights into language development trajectories (Chuey et al., 2021).

The urgency of this shift is underscored by the growing body of research demonstrating that online methods can produce results consistent with traditional in-person studies (Chuey et al., 2024). However, rigorous validation is essential to establish their credibility and ensure that they can be reliably used to draw meaningful conclusions (Chuey et al., 2021; Rhodes et al., 2022). By replicating established in-person paradigms in online settings, researchers can not only test the validity of these methods but also contribute to addressing the replicability crisis by strengthening the robustness and generalizability of findings in child language research.

1.3 Current study

In the present study, we aimed to replicate an in-person, eye-tracking speaker reliability paradigm designed by G&K in an online experiment to assess novel word learning in 4–6-year-old, English-speaking monolingual children. The original study compared the performance of monolingual children with that of Spanish-English-speaking bilingual children. However, in our study, we focused exclusively on determining the replicability of the task and tested only monolingual children. Additionally, G&K found no significant group difference in task performance. Thus,

testing only monolingual children should be sufficient to evaluate the task's validity and obtain meaningful results.

We aimed to determine whether the patterns observed in controlled lab environments hold true in online settings, which offer greater ecological validity and accessibility. To test this hypothesis, we utilized the same experimental paradigm as G&K, meticulously replicating the stimuli features, timing parameters, and procedural steps. This included the use of identical audio-visual materials, the same sequence of experiment phases, and consistent intervals between trials. By maintaining these methodological constants, we ensured that any observed differences or consistencies in word learning would be attributed to the change in the mode of delivery (offline vs. online) rather than to variations in the experimental design. In the study conducted by G&K, the authors found that children learned new words taught by both the reliable and unreliable speakers, despite their preference for the reliable speaker. We hypothesized that if a speaker reliability cue produces similar results across lab-based and web-based methodologies, then children should also learn new words from both reliable and unreliable speakers in a web-based online experiment.

2 Methods

2.1 Participants

Four- and 6-year-old English-speaking monolingual children were recruited online through social media posts, flyers posted around the city of Bloomington, IN, and flyers distributed at local events and establishments. Inclusionary criteria for all children included exposure to English from birth and no significant exposure to another language (defined as >5% weekly exposure). Exclusionary criteria for all children included hearing loss or history of hearing difficulties, psychological or behavioral disorders, neurological impairments, other developmental disabilities, and scoring 1.5 standard deviations below the mean on a non-verbal IQ test. Children were also excluded if they met any two of the following criteria: standardized vocabulary scores below 85, diagnosis of a language impairment, or parent concerns regarding their child's language development. A total of 23 children participated. Based on the abovementioned criteria, 3 children were excluded from the study (low IQ score, missing exposure data, and exposure to a second language). No children were excluded because of missing experimental or assessment data. Thus, the present study included a total of 20 monolingual children ($M_{\text{age}} = 5.33$, $SD_{\text{age}} = 1.02$). They were tested over two 1-h sessions (see Table 1 for participant characteristics). All data collection was conducted remotely via Zoom. Parents provided written consent for their and their child's participation, and all study procedures were approved by Indiana University's Institutional Review Board.

2.2 Parent questionnaires

All parents filled out a background questionnaire about the child's language, medical and educational histories. The form also yielded mother's years of education to index the child's socio-economic status (SES; Justice et al., 2020). Additionally, a history

TABLE 1 Participant characteristics.

	Mean (SD)
<i>N</i>	20 (10 females)
Age (years)	5.33 (1.02)
^a SES (years)	18.53 (2.80)
^b Non-verbal IQ (std. scores)	106.05 (16.66)
^c Vocabulary (std. scores)	110.68 (13.45)

^aSocio-economic status indexed by years of maternal education.

^bIndexed by the KBIT-2, Matrices subtest.

^cMeasured by the ROWPVT-4th edition.

of the child's language development and language exposure were obtained through a parent interview.

2.3 Assessments

All children were administered the Visual Matrices subtest of the Kaufman Brief Intelligence Test—Second Edition (KBIT-2; Kaufman and Kaufman, 2004), a test of non-verbal intelligence. To assess English vocabulary knowledge, children were given the Receptive One-Word Picture Vocabulary Test, 4th Edition (ROWPVT-4; Martin and Brownell, 2010).

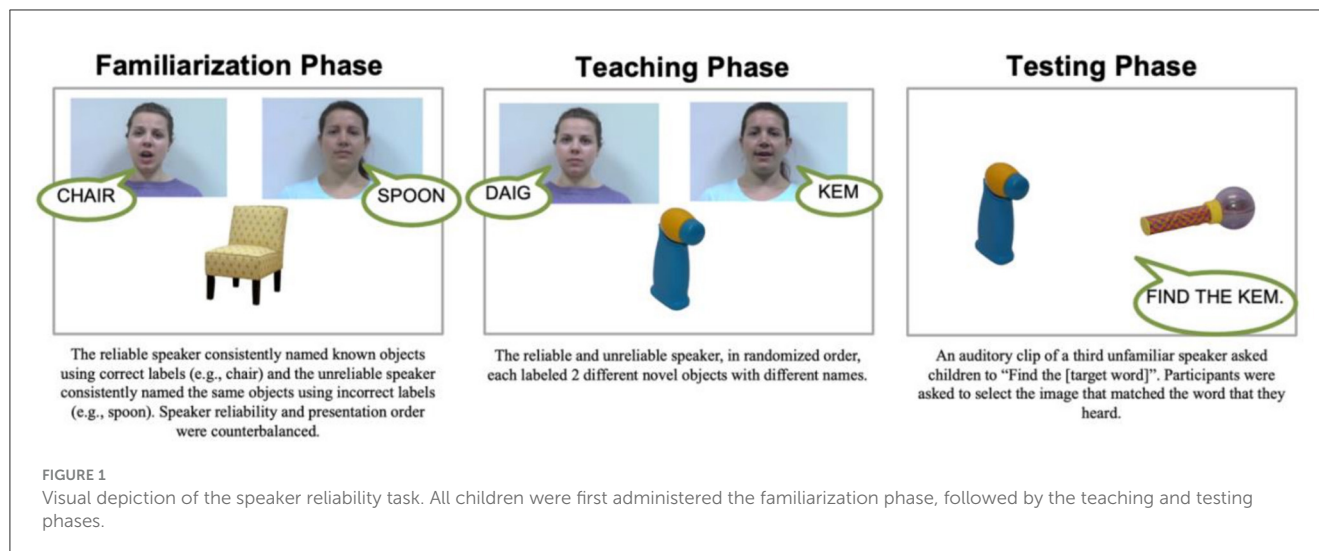
2.4 Experimental task

The task was hosted on Gorilla (<https://gorilla.sc>), a platform that allows researchers to build experiments online. Each child completed an online version of the speaker reliability task adapted from G&K, where all original stimuli and recordings were used for the online version of the experimental task. Children were seated in front of a computer screen in their homes. In the familiarization phase (see Figure 1 for a visual description of the entire task), 6 familiar objects were presented on the screen by two female speakers, one at a time, in random order. For each object, the first speaker appeared on the screen and labeled the object correctly (e.g., labeling “doll” as “doll”). Then, the second speaker appeared and labeled the same object incorrectly (e.g., labeling “doll” as “house”). Thus, one speaker consistently provided correct labels for the objects (the reliable speaker) and the other speaker consistently provided incorrect labels for the same objects (the unreliable speaker). The order (who appeared first on the screen) and the position (who appeared on the right side of the screen) of the speakers, as well as their reliability status (which speaker was reliable) were counterbalanced. Immediately after the familiarization trials, children were shown pictures of the two speakers on the screen and were asked the following questions.

1. Was one of them giving you all the right names? Which one?
2. Was one of them giving you all the wrong names? Which one?
3. Which one did you like better?
4. Why did you like her better?

No feedback was given to the children during these questions.

After familiarization, children were taught new words by the two speakers in a teaching phase. A single novel object was



presented on the screen and both speakers labeled the object one by one (e.g., the reliable speaker said “Kem” while the unreliable speaker said “Peem”). Children were taught novel labels for 2 novel objects by the reliable and unreliable speakers and were exposed to the novel words twice, in randomized order. Word-object pairings, the reliability of the speaker, the speaker position, and the speaker presentation order were counterbalanced across children. Then, in the testing phase, the two novel objects appeared on the screen, followed by the auditory carrier phrase, “Find the...Peem.” Children were given 10 s from word onset (e.g., “Peem”) to click on the picture that matched the word they heard. All testing trials only included pictures of the two novel objects; thus, the distractor object was always the other novel object and there were no additional, unfamiliar novel objects presented during testing. Each novel label taught by both speakers was tested 3 times, resulting in 12 testing trials. Position of the novel objects was counterbalanced across children, and trial presentation was randomized. Children’s accuracy of responses was collected.

All aspects of the original task in G&K, including stimuli and inter-stimulus times, were maintained, except for the following modifications made for online administration. First, children’s responses were collected via mouse clicks. Second, all testing trials were 10 s long, rather than 4 s, because piloting revealed that children required more time to use a mouse. Third, for questions 1–3 in the familiarization phase, children responded by clicking on the picture of one of the two speakers, instead of pointing to it.

2.5 Analyses

Data cleaning and analyses were carried out in RStudio 2024.04.2. Children’s accuracy was calculated as the proportion of correct responses (i.e., clicks on the target object). The target object was defined as the novel object that matched the novel label prompt. The novel label was either uttered previously by the reliable speaker or the unreliable speaker. Two types of analyses were conducted. First, Wilcoxon signed rank *t*-tests were conducted to determine whether children exceeded chance levels

of performance in both experimental conditions. Second, a logistic mixed-effects model was estimated in which children’s accuracy (0, 1) was regressed on Condition (contrast coded as -0.5 for Reliable and 0.5 for Unreliable). The final model included a fixed effect of Condition, along with random effects of participant and item.

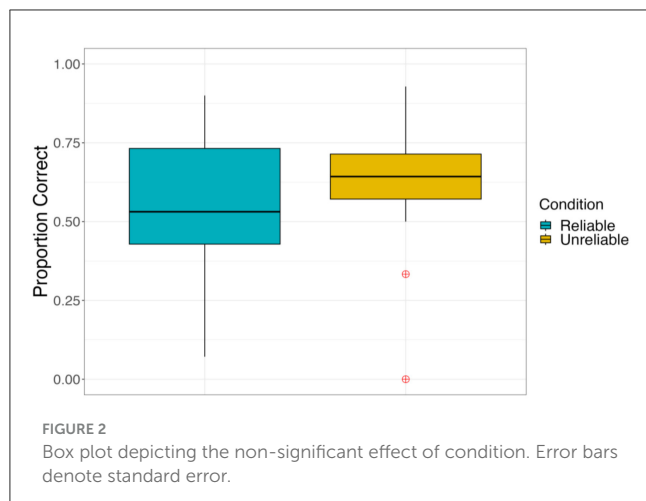
3 Results

3.1 Explicit probes of speaker reliability status

Responses to the questions about speaker reliability following the familiarization phase were examined. Seventeen children (out of 20) correctly identified the reliable speaker as always providing correct labels and the same children correctly identified the unreliable speaker as always providing incorrect labels. When asked who they liked better and why, 13 children said that they liked the reliable speaker better because “she gave the right answers,” 3 children said, “I don’t know,” 2 children did not respond, and 2 children selected the reliable speaker because of their physical description (e.g., “her hair was nice”). Because 3 children misidentified the speakers, we analyzed the data with and without them. The analyses yielded similar results with and without these children included. Therefore, below we present the analysis for all 20 children.

3.2 Mean accuracy

One sample *t*-tests showed that all children exceeded chance levels of performance in both the reliable ($M = 0.55$, $SD = 0.50$, $p = 0.02$) and unreliable ($M = 0.61$, $SD = 0.49$, $p = 0.005$) conditions. The logistic mixed-effects model revealed no main effect of Condition ($b = 0.06$, $t = 0.89$, $p = 0.34$), indicating that children’s task accuracies for the reliable and unreliable conditions were not significantly different (see Figure 2).



4 Discussion

This study aimed to replicate lab-based, eye-tracking findings reported in G&K, examining children's word learning from a speaker reliability cue in a fully online paradigm using mouse clicks. Results were consistent with those of G&K, showing that children performed above chance in each learning condition (reliable and unreliable), and that there was no main effect of condition. These findings validate the use of speaker reliability, as manipulated in this study, in an online setting as they demonstrate that children respond similarly to these cues whether tested in person or online. The robustness of this particular reliability manipulation is further strengthened by the use of the exact same stimuli as G&K. Replicating studies using identical stimuli across different experimental setups is typically challenging due to variations in participant demographics, environmental factors, and methodological differences (Open Science Collaboration, 2015). It should be noted that the G&K study included both monolingual and bilingual children. While the study reported no bilingual effects on reliability, it is entirely possible that group differences could emerge in an online setting. Thus, future work must include bilingual children to expand on this replicative effort. Nonetheless, by successfully using the same stimuli in an online format, this study contributes to addressing the replicability crisis in psychological research (Ioannidis, 2005). It demonstrates that children's responses to speaker reliability cues are stable and dependable across different methodologies. This finding not only validates the effectiveness of speaker reliability cues in word-learning studies but also highlights the potential for online paradigms to produce high-quality, replicable research outcomes.

However, the study faced several obstacles and limitations that are crucial for future online research to consider. The most frequent issues were technical breakdowns, including device compatibility and internet connectivity problems. These disruptions led to interruptions, resulting in data loss or incomplete sessions; children with missing data were consequently excluded from the analysis. While online data collection has the advantage of reaching a larger pool of participants, it also increases the potential for data loss due to technical difficulties. Addressing these challenges is essential to

enhance the effectiveness of online studies. We also encountered some variability in parental assistance. Although most parents appropriately supported their children during the sessions, some provided excessive or insufficient guidance. We anticipated this challenge and implemented strict protocols for parents, especially to minimize excessive guidance. These procedures ensured, as much as possible, the collection of reliable data. Online child studies should account for variability in parental input, as it can make it difficult to maintain research control, and it can make the results more difficult to interpret. Another challenge was to ensure the accuracy and integrity of data collected online, particularly with younger children who may not follow instructions as closely without in-person supervision. Thus, implementing rigorous protocols for children became essential, which included providing clear, age-appropriate instructions and incorporating interactive elements to engage children effectively. Finally, the study's sample size may not be fully representative of all children who participate in online research. While the sample size aligns with prior studies and is adequate for the study's goals, selection bias remains a concern. Families participating in research—particularly in online contexts—may have characteristics that differ from the broader population, such as better access to technology or the internet, which is especially problematic for disadvantaged learners. Although efforts were made to recruit a representative sample, this limitation highlights the importance of replicating studies across different data collection modes to capture diverse demographics.

Overall, while this study confirms the feasibility of conducting language research online with young children and replicates the findings of G&K in an online format, it also underscores the need to address several challenges inherent in online data collection. This study not only contributes to the growing body of evidence supporting online research but also provides valuable insights into optimizing online experimental designs for future investigations.

5 Conclusion

Embracing online methodologies is not just a response to current challenges but a proactive step toward future-proofing child language research. It opens up new possibilities for innovative study designs, increased participant engagement, and the potential to reach global populations. In an increasingly digital world, the ability to conduct high-quality, scalable research online is imperative for the continued advancement of our understanding of child language development. To this end, our study examining the effect of speaker reliability on word learning in children provides evidence that web-based, online studies are feasible and can yield results that are consistent with lab-based studies.

Future work in this area should extend the use of online paradigms to include more sophisticated measures (e.g., web-based eye-tracking) to study other aspects of word learning (e.g., looking patterns) and to examine other types of speaker cues. Future studies will also benefit from utilizing online platforms to include a more diverse range of participants, including children from various linguistic, cultural, and socio-economic backgrounds, to examine whether these factors widely influence performance on online language-learning tasks.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Institutional Review Board at Indiana University-Bloomington. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

IG: Conceptualization, Formal analysis, Funding acquisition, Methodology, Project administration, Software, Supervision, Writing – original draft, Writing – review & editing. LP: Data curation, Investigation, Methodology, Writing – original draft.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This project

was partially funded by the Vice Provost for Research through the Social Sciences Research Funding Program at Indiana University-Bloomington.

Acknowledgments

The authors would like to thank all of the families who participated in this study. We are grateful to the members of the Language Experience and Multilingualism Research Lab for their assistance with recruitment, data collection and data coding. Particular thanks to Hannah Clancy for programming the experiment.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Bascandziev, I., and Harris, P. L. (2016). The beautiful and the accurate: are children's selective trust decisions biased? *J. Exp. Child Psychol.* 152, 92–105. doi: 10.1016/j.jecp.2016.06.017
- Birch, S. A., Vauthier, S. A., and Bloom, P. (2008). Three- and four-year-olds spontaneously use others' past performance to guide their learning. *Cognition* 107, 1018–1034. doi: 10.1016/j.cognition.2007.12.008
- Chuey, A., Asaba, M., Bridgers, S., Carrillo, B., Dietz, G., Garcia, T., et al. (2021). Moderated online data-collection for developmental research: Methods and replications. *Front. Psychol.* 12:734398. doi: 10.3389/fpsyg.2021.734398
- Chuey, A., Boyce, V., Cao, A., and Frank, M. C. (2024). Conducting developmental research online vs. in-person: a meta-analysis. *Open Mind* 8, 795–808. doi: 10.1162/opmi_a_00147
- Corriveau, K., and Harris, P. L. (2009). Choosing your informant: weighing familiarity and recent accuracy. *Dev. Sci.* 12, 426–437. doi: 10.1111/j.1467-7687.2008.00792.x
- Corriveau, K. H., Meints, K., and Harris, P. L. (2009). Early tracking of informant accuracy and inaccuracy. *Br. J. Dev. Psychol.* 27, 331–342. doi: 10.1348/026151008X310229
- Cossette, I., Fobert, S. F., Slinger, M., and Brosseau-Liard, P. E. (2020). Individual differences in children's preferential learning from accurate speakers: Stable but fragile. *J. Cogn. Dev.* 21, 348–367. doi: 10.1080/15248372.2020.1727479
- Dautriche, I., Goupil, L., Smith, K., and Rabagliati, H. (2021). Knowing how you know: toddlers reevaluate words learned from an unreliable speaker. *Open Mind* 5, 1–19. doi: 10.1162/opmi_a_00038
- Feijoo, S., Amadó, A., Sidera, F., and Aguilar-Mediavilla, E. (2023). Language acquisition in a post-pandemic context: the impact of measures against COVID-19 on early language development. *Front. Psychol.* 14:1205294. doi: 10.3389/fpsyg.2023.1205294
- Fernald, A., Zangl, R., Portillo, A. L., and Marchman, V. A. (2006). "Looking while listening: using eye movements to monitor spoken language comprehension by infants and young children," in *Developmental Psycholinguistics: On-line Methods in Children's Language Processing* (John Benjamins Publishing Company), 113–132.
- Fusaro, M., and Harris, P. L. (2008). Children assess informant reliability using bystanders' non-verbal cues. *Dev. Sci.* 11, 771–777. doi: 10.1111/j.1467-7687.2008.00728.x
- Gangopadhyay, I., and Kaushanskaya, M. (2022). The effect of speaker reliability on word learning in monolingual and bilingual children. *Cogn. Dev.* 64:101252. doi: 10.1016/j.cogdev.2022.101252
- Hoffmann, S., Tschorn, M., and Spallek, J. (2024). Social inequalities in early childhood language development during the COVID-19 pandemic: a descriptive study with data from three consecutive school entry surveys in Germany. *Int. J. Equity Health* 23:2. doi: 10.1186/s12939-023-02079-y
- Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Med.* 2:e124. doi: 10.1371/journal.pmed.0020124
- Jaswal, V. K., and Neely, L. A. (2006). Adults don't always know best: preschoolers use past reliability over age when learning new words. *Psychol. Sci.* 17, 757–758. doi: 10.1111/j.1467-9280.2006.01778.x
- Justice, L. M., Jiang, H., Bates, R., and Koury, A. (2020). Language disparities related to maternal education emerge by two years in a low-income sample. *Matern. Child Health J.* 24, 1419–1427. doi: 10.1007/s10995-020-02973-9
- Kaufman, A. S., and Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test, Second Edition*. Bloomington, MN: Pearson, Inc.
- Kinzler, K. D., Corriveau, K. H., and Harris, P. L. (2011). Children's selective trust in native-accented speakers. *Dev. Sci.* 14, 106–111. doi: 10.1111/j.1467-7687.2010.00965.x
- Koenig, M. A., Clément, F., and Harris, P. L. (2004). Trust in testimony: children's use of true and false statements. *Psychol. Sci.* 15, 694–698. doi: 10.1111/j.0956-7976.2004.00742.x

- Koenig, M. A., and Harris, P. L. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Dev.* 76, 1261–1277. doi: 10.1111/j.1467-8624.2005.00849.x
- Koenig, M. A., and Woodward, A. L. (2010). Sensitivity of 24-month-olds to the prior inaccuracy of the source: possible mechanisms. *Dev. Psychol.* 46:815. doi: 10.1037/a0019664
- Krogh-Jespersen, S., and Echols, C. H. (2012). The influence of speaker reliability on first vs. second label learning. *Child Dev.* 83, 581–590. doi: 10.1111/j.1467-8624.2011.01713.x
- Lefever, S., Dal, M., and Matthíasdóttir, Á. (2007). Online data collection in academic research: advantages and limitations. *Br. J. Educ. Technol.* 38, 574–582. doi: 10.1111/j.1467-8535.2006.00638.x
- Liu, D., Vanderbilt, K. E., and Heyman, G. D. (2013). Selective trust: children's use of intention and outcome of past testimony. *Dev. Psychol.* 49:439. doi: 10.1037/a0031615
- Mangardich, H., and Sabbagh, M. A. (2018). Children remember words from ignorant speakers but do not attach meaning: evidence from event-related potentials. *Dev. Sci.* 21:e12544. doi: 10.1111/desc.12544
- Martin, N. A., and Brownell, R. (2010). *Receptive One-word Picture Vocabulary Test Fourth Edition (ROWPVT-4)*. Novato: Academic Therapy Publications.
- Nurmsoo, E., and Robinson, E. J. (2009). Identifying unreliable informants: do children excuse past inaccuracy? *Dev. Sci.* 12, 41–47. doi: 10.1111/j.1467-7687.2008.00750.x
- Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science* 349:aac4716. doi: 10.1126/science.aac4716
- Pasquini, E. S., Corriveau, K. H., Koenig, M., and Harris, P. L. (2007). Preschoolers monitor the relative accuracy of informants. *Dev. Psychol.* 43, 1216–1226. doi: 10.1037/0012-1649.43.5.1216
- Rhodes, M., Leslie, S.-J., and Tworek, C. M. (2022). Validity and reliability in online child language research. *Dev. Sci.* 25:e13105.
- Rhodes, M., Rizzo, M. T., Foster-Hanson, E., Moty, K., Leshin, R. A., Wang, M., et al. (2020). Advancing developmental science via unmoderated remote research with children. *J. Cogn. Dev.* 21, 477–493. doi: 10.1080/15248372.2020.1797751
- Sabbagh, M. A., and Shafman, D. (2009). How children block learning from ignorant speakers. *Cognition* 112, 415–422. doi: 10.1016/j.cognition.2009.06.005
- Scofield, J., and Behrend, D. A. (2008). Learning words from reliable and unreliable speakers. *Cogn. Dev.* 23, 278–290. doi: 10.1016/j.cogdev.2008.01.003
- Sobel, D. M., and Macris, D. M. (2013). Children's understanding of speaker reliability between lexical and syntactic knowledge. *Dev. Psychol.* 49, 523–532. doi: 10.1037/a0029658
- Sobel, D. M., Sedivy, J., Buchanan, D. W., and Hennessy, R. (2012). Speaker reliability in preschoolers' inferences about the meanings of novel words. *J. Child Lang.* 39, 90–104. doi: 10.1017/S0305000911000018
- Tomasello, M. (2020). Evolution of human communication and language. *Curr. Opin. Psychol.* 32, 146–151. doi: 10.1016/j.copsyc.2019.08.014
- Wood, L. A., Kendal, R. L., and Flynn, E. G. (2013). Whom do children copy? Model-based biases in social learning. *Dev. Rev.* 33, 341–356. doi: 10.1016/j.dr.2013.08.002