Advancing Developmental Science via Unmoderated Remote Research with Children

Marjorie Rhodes , Michael T. Rizzo , Emily Foster-Hanson , Kelsey Moty , Rachel A. Leshin , Michelle Wang , Josie Benitez & John Daryl Ocampo

Published online: 13 Aug 2020.

Submit your article to this journal ☐

View related articles ☐

View Crossmark data ☐

Routledge
Taylor & Francis Group

Advancing Developmental Science via Unmoderated Remote Research with Children

Marjorie Rhodes[a], Michael T. Rizzo[a,b], Emily Foster-Hanson[a], Kelsey Moty[a], Rachel A. Leshin[a], Michelle Wang[a], Josie Benitez[a], and John Daryl Ocampo[a]

[a]New York University; [b]Beyond Conflict Innovation Lab, Boston, MA

ABSTRACT

This article introduces an accessible approach to implementing unmoderated remote research in developmental science – research in which children and families participate in studies remotely and independently, without directly interacting with researchers. Unmoderated remote research has the potential to strengthen developmental science by: (1) facilitating the implementation of studies that are easily replicable, (2) allowing for new approaches to longitudinal studies and studies of parent-child interaction, and (3) including families from more diverse backgrounds and children growing up in more diverse environments in research. We describe an approach we have used to design and implement unmoderated remote research that is accessible to researchers with limited programming expertise, and we describe the resources we have made available on a new website (discoveriesonline.org) to help researchers get started with implementing this approach. We discuss the potential of this method for developmental science and highlight some challenges still to be overcome to harness the power of unmoderated remote research for advancing the field.

The field of cognitive development was founded upon remarkable insights gleaned from everyday interactions with children. Piaget's theory of cognitive development (1954) began with his observations of his own children playing on their mats, dropping things from their highchairs, and playing with marbles. Carey (1985) revolutionized our understanding of how concepts originate and change by analyzing conversations with her own child about birth, the nature of life, and death. The field is full of stories of great theoretical insights made by researchers closely watching children as they crawl around near the sides of high beds (Adolph, Kretch, & LoBue, 2014), negotiate the rules of games among themselves on a playground (Borman, 1981), and try to sit down on way-too-tiny toy tractors (DeLoache, 1987). Of course, the field has never relied on the observations of individual researchers alone. We use these observations to design experiments that recreate the situations in which the behaviors were first observed, which can then be replicated by labs around the world. But still, individual interactions between researchers and their participants have always been central to the field of cognitive development. The notion that we can learn from watching and interacting with children, along with the intriguing challenges of thinking about how to recreate the conditions of everyday life in

the lab to elicit behaviors of interest and test their underlying mechanisms, draws many students to the field.

And yet, this approach has always had limitations. First, there is the thorny issue of Clever Hans[1] – that researchers interacting directly with their participants might bias them toward particular responses, especially when researchers know the hypotheses and theoretical issues at stake. Second, there is the issue that complex behavioral paradigms can be difficult for other labs to replicate – allowing disagreements to develop over whether different findings reflect slight variations in the testing protocols, the skill of the student or research assistant, or other interpersonal or contextual factors – making cross-lab replication challenging (for discussion, see Adolph, Gilmore, & Kennedy, 2017; Gilmore, Kennedy, & Adolph, 2018). Third, these studies are very labor intensive both in terms of recruitment (finding enough families willing and able to visit labs or other testing locations) and administration (often involving multiple highly-trained experimenters), perhaps contributing to small sample sizes and limitations with statistical power (for problems with small sample sizes, see Button et al., 2013). And fourth, the need for researchers (or their staff or students) to interact with their participants directly often limits studies geographically. Though we as developmental researchers are interested broadly in *children*, we often study those within driving distance of our labs. This can contribute to a lack of diversity – with respect to race, ethnicity, religion, economic background, and other aspects of social experience – in our samples (Nielsen, Haun, Kärtner, & Legare, 2017). To compound the general issues described above, at the time we are writing this article, in-person research at many universities around the world has been suspended due to the global pandemic of Covid-19, making the reliance of our field on studies of in-person interaction particularly challenging.

These broad concerns are certainly not limitations of all in-person studies in the field of cognitive development, and there are many well-established and exciting new ways to mitigate them, including having blind experimenters, developing multi-lab collaborations, adopting new practices for data and protocol sharing, and making efforts to diversify samples for in-person research (see Chouinard, Scott, & Cusack, 2019; Gilmore et al., 2018; Rubio-Fernández, 2019; Scott, Chu, & Schultz, 2017; Scott & Schultz, 2017; The Many Babies Consortium, 2020). We note these limitations simply to acknowledge the challenges inherent to doing work in cognitive development – a science that, perhaps more than most other subdisciplines of psychology, has relied heavily on interpersonal interaction between the researcher and participant. Here we offer an additional tool for the developmental scientist's toolbox to strengthen our science as we move forward as a field.

This approach is unmoderated remote research with children. This simply means designing studies that children, on their own or with minimal assistance from their parents, can complete remotely, with data recorded by widely available online experimental testing software and the families' webcam. We call this approach *unmoderated* because children and families do not directly interact with researchers at all. This differentiates the approach that we are describing from those that rely on videoconferencing with an experimenter (for discussion of the strengths and challenges of video-conferencing approaches, see Sheskin & Keil, 2018).

---

[1]Clever Hans was a horse who could supposedly solve problems of multiplication and division through tapping (Pfungst, 1911). The Clever Hans Phenomenon describes the subtle ways in which experimenters can influence the responses of human and non-human research participants, often without realizing that they are doing so (Rosenthal, 1966).

We are not the first to use unmoderated remote research for developmental science (see Scott & Schultz, 2017; Scott et al., 2017 for discussion of the Lookit platform). To complement the growing number of approaches to this new general methodology, we add methods for presenting stimuli and collecting data from verbal children (ages 3 and older, rather than infants) and strategies for collecting and coding a wide variety of responses from webcam videos and online testing software (including language, non-verbal behaviors, button presses, computer clicks, timing information, and so on). Our goal in this article, and in a companion website we have developed (discoveriesonline.org), is to offer a practical guide for how to do such research with limited programming expertise. In the subsequent sections, we discuss how we collect informed consent; how we design and implement studies; how we process, code, and share data; and how we recruit families to participate. We discuss the potential and challenges of this new general method, and we describe two sample studies to illustrate our approach.

Overview of unmoderated remote research

Privacy and informed consent

In our approach to unmoderated remote research, we collect informed consent by video. After selecting a study that they are interested in, parents give permission for their webcam to activate and then click through a series of slides that provide standard informed consent information (see discoveriesonline.org/irb). Parents provide verbal consent for their child to participate by repeating a standardized consent statement (which is recorded). Children are next asked to provide verbal assent (also recorded). After the consent procedure, the study launches in the browser. The child and parent are recorded while they complete the study, and the experimental testing software (e.g., Qualtrics) that presents the study also records data (e.g., button presses, answer choices, response times, and so on). At the end of the study, parents set privacy settings for their video (including whether to make the video available only to the lab or to allow it to be shared securely with other researchers in the field via the digital data library, Databrary, 2012; Gilmore et al., 2018), and then upload the video to the research platform. If something unexpected happened in the video (e.g., it inadvertently recorded a family interaction that the parent would prefer not to share), parents can simply choose to not upload their video at the end.

Procedures for video storage and security are described during the informed consent process to parents and in more detail online (discoveriesonline.org/irb; and on our testing website itself, here: discoveriesinaction.org/accounts/privacy_policy/). In brief, videos are initially stored on a secure web server accessible only to associated research staff, and are then downloaded, erased from the web server, and stored on a university-supported secure server. We provide our sample IRB and consent language on discoveriesonline.org/irb, although universities vary in their concerns about and standards for remote research (and labs outside the United States may face additional privacy and data-sharing regulations that need to be addressed).

Study design and implementation

Studies well-suited for unmoderated remote research

Studies that involve telling children simple stories or prompts and asking them to select between pictured response options can be easily designed for unmoderated remote research

(see Figure 1). We provide further examples of sample stimuli, including those adapted from in-person published research, on discoveriesonline.org, as well as a detailed step-by-step guide on how to create similar stimuli and program these studies.

There are some additional kinds of developmental tasks and research questions that are particularly well-suited for an unmoderated remote approach. For example, remote testing allows researchers to code parent-child interactions in a home setting, reducing some barriers to participation (e.g., being within driving distance of the lab and having time to come in). Further, although parent-child research can be performed in a lab, or by sending a research assistant to videotape at participants' homes (e.g., Tamis-LeMonda, Custode, Kuchirko, Escobar, & Lo, 2019), unmoderated remote research allows families to participate in a more natural setting, without the presence of a stranger. From webcam data, researchers can code for parent and child language, gesture, facial expression and so on, while parents and children sit together at their home computers (Figure 2).

Unmoderated remote testing also reduces some barriers to longitudinal studies. Because families can participate from home, unmoderated remote research provides a straightforward opportunity for longitudinal studies of parent-child interaction, can track children and families even as they move or change schools, and does not require coordination of multiple lab or site visits. In our longitudinal studies, we simply notify families by e-mail when it is time for their next session, and families can complete the session of data collection within a given window when ever it is convenient for them (for an example, see https://osf.io/s36ek/).

There are, however, some types of studies that may not be well-suited for this approach. For instance, studies that involve behaviors that cannot be captured via webcam or require creating situations that cannot be animated might not be possible to conduct remotely. As examples, studies that examine children's helping behavior (e.g., Foster-Hanson, Cimpian, Leshin, & Rhodes, 2020; Warneken & Tomasello, 2006), crawling or walking over a novel apparatus (e.g., Adolph et al., 2003), searching or exploration behavior (e.g., Wu, Muentener, & Schulz, 2017), imitation behavior (e.g., Meltzoff & Moore, 1977), or social interactions with researchers (e.g., Hamann, Bender, & Tomasello, 2014; Martin & Fabes,

Figure 1. Items from an unmoderated remote study of children's gender beliefs (see https://osf.io/htn7z/for more information about the study). For example, children are asked: (1; *left*) What a baby will be like when it grows up after hearing that it was born a girl but raised entirely with boys, (2; *center*) What they themselves will be like when they grow up, and (3; *right*) Why a particular boy or girl has a novel property. Each question is presented as a short narrated video, with answer choices animated to pop or jiggle as a narrator names each choice. The order of answer choices on the screen is counterbalanced. Children click on the pictures to indicate their response, or they point and have a parent click for them if they are unable to click on their own.
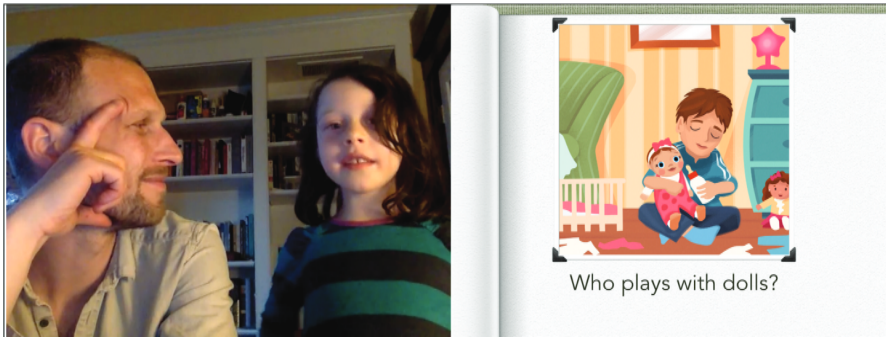
Figure 2. A parent and child completing an unmoderated remote study on parent-child conversations about gender. Parents and children look through pictures on the screen showing gender-typical and -atypical behaviors and are asked to discuss them. We then code their conversations from webcam video using Datavyu (see https://osf.io/htn7z/for more information about the study).

2001) may be more difficult to conduct with this remote approach, although, under some circumstances, could still be feasible.

Designing effective studies for unmoderated remote research

One possible concern about unmoderated studies is that children will lose interest, and without an experimenter present to reengage them, data quality and retention can drop off. Although some studies are indeed better suited to a live experimenter, we have found that simple animations can go a long way toward creating child-friendly, unmoderated studies that children find engaging and can complete on their own, with little help (or unwanted interference) from their parents. In the same way that an experimenter may point to a picture or present stimuli one at a time, animations can highlight important elements in a task. For instance, for the measures presented in Figure 1, as the recorded audio states each option aloud, the corresponding button on the screen animates (e.g., the green circle and red circle in Figure 1, center, jiggle when the narrator asks aloud, "yes" or "no"). Animations can also help maintain children's attention throughout the task; for example, we have observed that changing the narrator's physical gestures (e.g., having the narrator point to the child when it is their turn to respond) or facial expressions (e.g., changing the narrator's expression from smiling to inquisitive when asking a question), as well as highlighting important stimuli (e.g., with a small pulse or jiggle), can help maintain attention throughout. Programming contingent or corrective feedback into studies is also straightforward and can help to both document comprehension and reinforce critical manipulations. In these ways, animations can mirror the kind of feedback that children typically receive from in-person experimenters, and in a manner that is more consistent across participants throughout a study.

In unmoderated remote research, a number of issues (e.g., the quality of participants' internet connection) are outside of researchers' control, but there are several steps researchers can take to help studies run smoothly. For instance, we find it helpful to minimize how often children need to "click" to move through the study by using auto-advancing functions that help children progress to the next screen. Also, all video and audio files should be compressed to optimize them for web display; this will help reduce

the load on participants' internet connection and subsequent lags in displaying stimuli. Finally, especially for longer studies, researchers should consider providing a button within the study that allows parents to skip to the end of the study if their child no longer wishes to participate. In addition to reinforcing that participation is voluntary, this also allows for the possibility of collecting partial data that might otherwise be lost. A guide for creating animations (using Keynote or PowerPoint), recording audio using widely available software, designing child-friendly response options, and programming studies that will run in a child-friendly manner (in Qualtrics) is available at discoveriesonline.org/design.

Discoveries Online

As part of our system for unmoderated remote research for our lab, we developed an online platform for presenting studies to families via their internet browsers, recording the webcam data, and uploading the videos securely. In our own lab, we have integrated this platform for data collection within a larger system for participant recruitment and management. While the participant database and management systems are unique to our lab, the component of the platform that presents the study, captures video and test data, and securely stores video data, which we call Discoveries Online, can host studies for other labs as well, and may present a good option for labs who are first trying out this approach and do not wish to create their own video-recording and storage platform (see discoveriesonline.org/hostmystudy).

Data processing, coding, and sharing

Initial data processing

After parents have successfully uploaded a video recording, the first steps of data processing include making sure that (1) a child and parent were present during the study, (2) parent consent and child assent were clearly provided, and (3) the video is long enough to capture the study session. Other criteria may also be added to these initial checks for particular studies. For example, a study assessing parent-child conversations may require that a video have high-quality audio, that the parent and child interact for a minimum amount of time, or that parent-child dyads have their faces in full view of the camera.

Interference and the role of parents

One key difference between remote and in-person research is the testing environment, and who is (and is not) present during study administration. Specifically, in remote research, a researcher is not present, but the child's parent is, along with siblings and anyone else in the home. With unmoderated remote research, families never have direct contact with a researcher. This difference in the testing environment presents both unique challenges and opportunities.

One main challenge in remote research is ensuring that children's responses are not influenced by interference from other people. We address this issue in two ways: (1) by designing studies that are age-appropriate so that children can complete them with minimal assistance from parents, and (2) by coding video data for evidence of interference.

The first step in minimizing interference is making sure that studies are designed so that children can complete them as independently as possible (see Figure 1). For studies

intended to examine how children respond with as little interference from parents as possible, we include a note to parents at the start of the study asking them to allow the child to complete the study on their own:

> There are no right or wrong answers; we are just interested in what children think. If your child is able to use a mouse, please allow them to do the clicking. If your child is unfamiliar with using a mouse, they can point to the screen to indicate their answers and you can do the clicking for them. Please do not provide your child with any feedback throughout the study – we are interested in what children think on their own.

Second, we code participants' videos for interference using Datavyu, a free open-source video coding software (Datavyu.org; Lingeman, Freeman, & Adolph, 2014). Interference coding involves having a trained researcher watch each video and mark any portions of the study in which a parent or other person interfered with the child's response (for detailed procedures, see discoviesiesonline.org/datavyu). To code for interference, the study has to be set up to allow us to track what participants are seeing at any given time during the video (since, in our system, the screen that participants see is not recorded). One way of doing so is to embed timing data in the experimental design software to record when each trial begins and ends, which we can then import directly into Datavyu. This way, we can accurately link what we see participants doing in their webcam videos (e.g., receiving feedback on a question from parents) with what participants were seeing on their computer screens throughout the study, and remove trials with interference as appropriate.

The proportion of videos that we code varies by study, depending on both the topic and the study design. For studies in which the likelihood of parental interference is high, or in which the presence or content of interactions between children and their parents is a dependent variable (e.g., when measuring conversations between parents and children), we code 100% of videos. For other studies, in which the likelihood of parental involvement is low and not of interest in itself, researchers may code a randomly selected subset of videos within each condition in order to estimate the interference rate. Our previous unmoderated remote research has identified extremely low rates of parent or sibling interference (less than 1% of trials) when parents are given instructions similar to those above, and excluding such trials has not influenced the overall pattern of results (Leshin, Leslie, & Rhodes, 2020). Based on these initial studies (which involved coding all trials), our current procedure for studies where we expect little interference is to randomly select 20% of videos in each condition to be coded by a trained researcher in order to estimate the interference rate. If the interference rate in the coded videos is lower than 3%, we conclude that the interference level is negligible and highly unlikely to influence the study patterns, and therefore retain all trials for analyses. However, if the interference rate in the subset of coded videos is more than 3%, we code every video trial-by-trial for instances of parental interference (with 20% also checked by a reliability coder) and drop all questions with interference from the analyses. In these cases, we also plan to drop participants if more than 25% of their responses contain potential interference (in practice, however, this has occurred very rarely). All plans for interference coding and drop criteria are determined in advance and pre-registered (see discoveriesonline.org/pre-registration).

As noted above, unmoderated remote research also provides unique opportunities to observe the behavior of children and their parents in a more naturalistic home setting,

reducing concerns about self-presentation and experimenter demands (Dahl, 2017; Engelmann, Herrmann, & Tomasello, 2012). In studies where parent-child interactions are the focus, unmoderated remote research allows for an analysis of the linguistic features of parent-child conversations as well as various behavioral cues that emerge during parent-child interactions (e.g., directional or communicative looking, pointing, shrugging, etc.; see Figure 2). For example, in one pilot study, we provided parents and children with an online version of a popular children's book that touched on issues related to interracial friendships and inclusion. Parents were asked to simply read and discuss the book with their children as they normally would. Videos were then coded and transcribed to examine differences in how parents talked – or did not talk – to their children about racial diversity and inclusion. Whereas getting parents and children to come to a lab setting to participate in a study about parent-child conversations can be expensive and labor-intensive, conducting these studies remotely is no more costly or time-consuming than any other type of study, since parents are already present throughout the study.

Sample recruitment and demographics

There are several ways to recruit families for unmoderated remote research. For instance, researchers might recruit participants for specific studies or try to build a database of families who are interested in participating in various studies over time. Here we describe how we recruited families to a database the characteristics of our sample to date, so that others can consider the strengths and limitations of various recruitment methods.

We relied heavily on popular social media platforms (e.g., Facebook, Instagram, Reddit, Twitter), including both maintaining active accounts and launching paid advertising campaigns on these sites. We also purchased advertisements on popular parenting podcasts and blogs. To supplement these recruitment efforts, we conducted in-person recruitment in local museums, schools, and community programs. For examples of our social media efforts, see discoveriesonline.org/recruitment.

Whereas we opted to create a database of interested families, very similar methods could be used to recruit for particular studies. Labs might find this to be a more desirable approach if they want to target recruitment from specific populations for particular studies, or as an initial way to try out this new approach. To recruit for particular studies, labs can post their studies on landing pages for online developmental studies (including childrenhelpingscience.com) and e-mail available participant databases (e.g., those they draw on for in-person research) with links to the study. Labs could also create separate, external landing pages that explain the study process and then link to a study hosted on an unmoderated remote research platform before starting the study. Further, it is possible to post single unmoderated remote studies via research recruitment platforms, such as Prolific.

We have conducted three cross-sectional studies and one longitudinal study drawing on the database of families that we recruited since launching our initial efforts with unmoderated remote research. Across these studies, a total of 1021 children have participated (51% female, 49% male) in the first 15 months that we operated the online lab, from across the United States (information on those who participated from other countries is below). The average age of participants was 6.30 years old (*SD* = 1.58, range 3.06 to 12.96 years). Racial/ethnic demographics for each child were provided by 99.8% of parents when registering their child in our database: 5.70% were Asian, 4.52% were Black (including 0.9% Black-

Hispanic), 4.43% were Latinx, 72.85% were White (including 2.26% White-Hispanic), 10.95% were bi- or multiracial, and 0.23% listed their race as "Other". Language information for each child was provided by 83.25% of parents; of those participants, 80.35% were monolingual English speakers, 19.53% were bi- or multilingual (including English), and 0.12% did not speak English. As noted above, at the end of each research session, parents were asked for their consent to upload their child's video to Databrary and for permission to use clips of the video in academic talks. Approximately 37% of parents provided consent to upload their videos to Databrary and 25% gave permission to use clips of their video in academic talks.

We gathered additional demographic information from some parents (those with children participating in studies interested in parental beliefs) to further analyze the features of our sample recruited to date. Of the 647 families we surveyed, 425 (66%) provided information about their political beliefs; a subsample of 136 (out of 188 asked, 72%) also provided information about their education level. Most of these parents were liberal and held at least a bachelor's degree; however, we found significant variation on both of these metrics (see Figure 3).

To assess the geographic reach of our recruitment efforts, zip code information was collected from parents upon registering for our database. Families came from 344 unique zip codes across 43 different states and the District of Columbia (see Figure 4).

Zip-code data were further analyzed by pulling 2019 census data for each of the 344 unique U.S. zip codes in our sample (census data were collected using the tool developed by Rizzo, Britton, Jamieson, & Rhodes, 2020, for identifying neighborhood-level variation in racial inequalities across the United States; https://osf.io/wybma/). Using these data, we examined participants' neighborhood type (Urbanized Area: 50,000 or more people, Urban Cluster: between 2,500 and 50,000 people, Rural: less than 2,500 people; classifications defined by the 2019 U.S. Census), racial diversity (proportion of racial/ethnic minority residents), median income, and education (proportion of residents completing at least some college; see Figure 5).
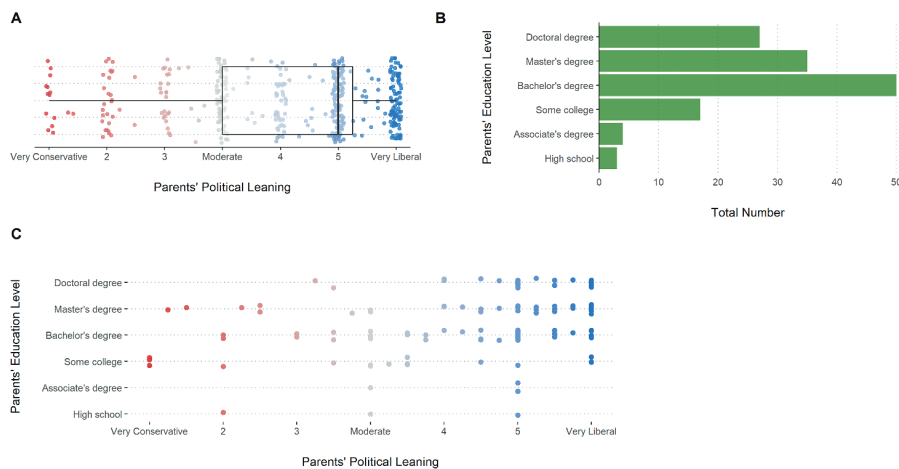

Figure 3. (a) Distribution of parents' political beliefs (N = 425). (b) Distribution of parents' education levels (N = 136). (c) The spread of education level by political leaning (N = 136).
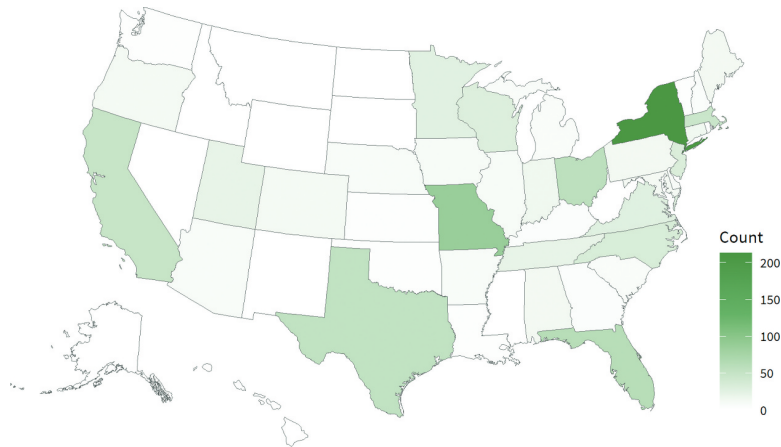
Figure 4. Regional diversity of U.S. participants. The fill of each state corresponds to the number of participants from that state (brighter green represents more participants).
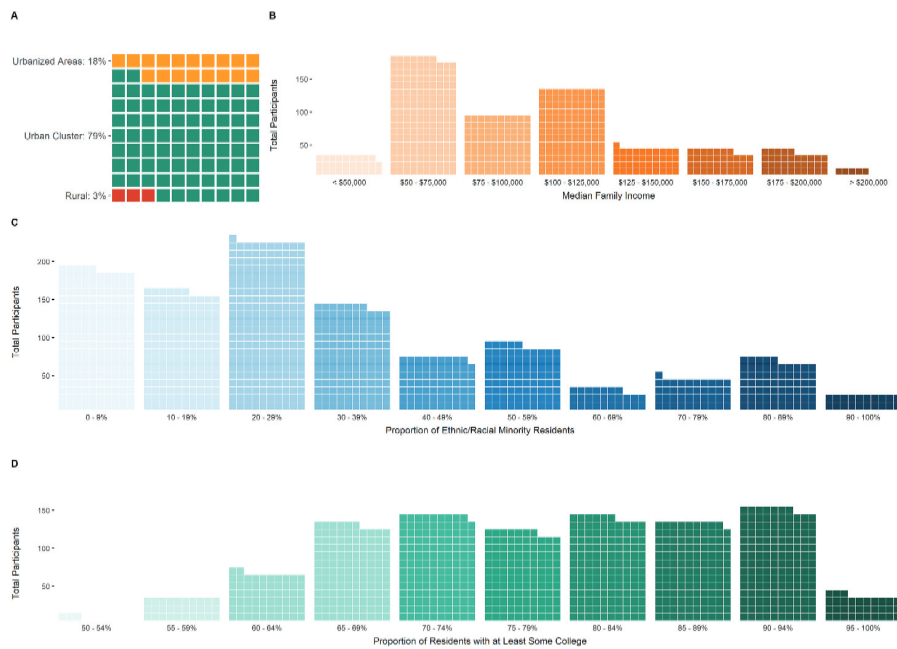

Figure 5. Neighborhood-level demographics for children who have participated in our online studies. (a) Proportion of participants living in each neighborhood type: rural, urbanized clusters, and urbanized areas. (b) Median family income for participant's zip code. (c) The proportion of ethnic/racial minority residents for participants' zip code. (d) The proportion of residents with at least some college for participants' zip code. Squares each reflect one participant, except in (a) where each square reflects 1% of participants.

We present these metrics here because they reflect demographic characteristics that may be of general interest to researchers within developmental science. Yet, more specific environmental metrics can also be examined based on the researcher's interests. For

example, in ongoing studies, we plan to use the tool developed by Rizzo et al. (2020) to examine how children's racial biases are related to environmental variation in racial inequalities (e.g., racial disparities in income, wealth, education, access to healthcare, and home ownership), political governance (e.g., results of local, state, and national elections), and implicit racial attitudes (assessed via Project Implicit; Xu, Nosek, & Greenwald, 2014). Ultimately, any variable of interest that can be tracked on a geographic level could be used to predict variation in children's development.

Unmoderated remote research has the potential to reach participants who are not typically included in developmental research. Henrich, Heine, and Norenzayan (2010) noted that the majority of children in developmental studies come from major urbanized areas around universities (as noted earlier, major urbanized areas are those with populations of greater than 50,000). In contrast, children in our unmoderated remote research studies have come from communities with more variability in population size, from a range of states and regions, and from neighborhoods with considerable variability in the demographic composition of children's neighborhoods, particularly with respect to race and ethnicity (Figures 4 and 5). Many aspects of cognitive development vary with respect to these and similar characteristics of children's local environments, including the development of folk-biological thought (Busch, Watson-Jones, & Legare, 2018; Medin, Waxman, Woodring, & Washinawatok, 2010; Ross, Medin, Coley, & Atran, 2003), essentialist beliefs about race and ethnicity (Mandalaywala, 2020), and theory of mind (Liu, Wellman, Tardif, & Sabbagh, 2008; Sabbagh, Xu, Carlson, Moses, & Lee, 2006). Thus, including children from more diverse communities (in terms of geography, demographics, and other aspects of children's social and physical environment) can improve the generalizability of study findings and also facilitate asking new questions about how various features of children's local environments shape cognitive development.

Currently, our database skews toward White, upper-middle class, and liberal families, relative to the U.S. population. Although we advertise broadly, our sample at this point still reflects a convenience sample of families who found our platform (for instance, the high levels of parental education in our sample probably reflect our tendency to talk about our unmoderated remote research at academic conferences and via Twitter). Yet, we think that remote research has great potential for improving diversity among participants in developmental science, particularly with respect to the inclusion of children from more diverse racial and ethnic backgrounds, but figuring out the best way to realize this potential requires additional work. For example, we are currently ramping up our recruitment efforts in more racially diverse school districts, exploring more targeted online media, and consulting with community partners to develop new ways to increase the diversity of our sample further.

One recruitment issue that is particularly challenging to overcome relates to economic diversity. Because unmoderated remote research – as we have implemented it – requires a computer with a webcam and reliable internet, participation is easiest for those who have access to this equipment in their homes. Families can participate from computers in other locations as well (e.g., schools, libraries, YMCAs, and so on), but these requirements do create a barrier for families without easy access to this technology (Lourenco & Tasimi, 2020). We hope to explore solutions to this challenge in our future work (for instance, by creating ways for families to participate via a broader range of technology), but for now this remains a challenge.

Using unmoderated remote research to facilitate robust, replicable, and open science

Unmoderated remote research provides an accessible tool for facilitating robust, replicable, and open science. As noted above, unmoderated remote research has the potential to increase the generalizability of research findings, by reducing some barriers to including larger and more diverse samples in developmental research. Unmoderated remote research can also improve replicability. In-person research often relies on complex interpersonal interactions, which can be difficult for other labs to fully reproduce, as well as for researchers within a single lab to reproduce with complete fidelity. In unmoderated remote research, everything has to be programmed and delivered without an experimenter present; therefore, the stimuli are easily shareable across labs, there is no way for experimenter expectations to influence the child's behavior, and complete consistency is maintained across administrations. In short, unmoderated remote research builds in a high level of experimental control (for general benefits of remote research for replicability and open science, see Gureckis et al., 2016). For these reasons, unmoderated remote research can facilitate cross-lab replication and attempts to build more closely on one another's science.

To further these aims, examples of pre-registration for some of our unmoderated remote research can be found at https://osf.io/8e2np/and https://osf.io/htn7z. These specify some details specific to unmoderated remote testing. For example, we pre-register details about data collection procedures (e.g., methods for participant recruitment and compensation), stopping rules (e.g., intentions to over-sample to account for video upload failures), data exclusion policies (e.g., guidelines for excluding trials with parental interference), and missing data practices (e.g., handling of partial data). To further facilitate potential cross-lab replication, using an entirely digital experimental protocol makes it straightforward to upload all stimuli, materials, and scripts onto the Open Science Framework (for example studies, see https://osf.io/acrwq/and https://osf.io/s36ek/). While all of these steps to promote robust and replicable science are also important for other types of research, we highlight them here because they are particularly straightforward to streamline into a workflow process for unmoderated remote research.

Examples of completed and ongoing studies

We completed our first study using unmoderated remote testing in the fall of 2019 (Leshin et al., 2020). Although this project sought to test novel empirical questions, one key goal was to conceptually replicate findings from previous work in the lab as a proof of concept that unmoderated remote research can capture the same developmental phenomena as more traditional in-person methods. We attempted to replicate a paradigm from Rhodes, Leslie, and Tworek (2012), in which children heard generic or specific language about a category, "Zarpies" (e.g., "*Zarpies* draw stripes on their knees" versus "*This Zarpie* draws stripes on her knees"), and then completed a series of measures probing essentialist beliefs about the category. The original study found that generic language increased children's essentialist beliefs – a finding that has broad implications for conceptual development, social cognition, and inter-group relations (e.g., Cimpian & Salomon, 2014; Moty & Rhodes, in press; Noyes & Keil, 2019; Rhodes, Leslie, Bianchi, & Chalik, 2018a; Rhodes, Leslie, Saunders, Dunham, & Cimpian, 2018b).

Stimuli were drawn from the original study and adapted for unmoderated remote research using the strategies outlined above (and described in detail on discoveriesonline.org/design).

Detailed descriptions of the methodology and procedures, including the ways in which our paradigm deviated from the original, were pre-registered at https://osf.io/acrwq/. In addition to testing the efficacy of unmoderated remote approaches to developmental research – particularly for questions that involve subtle linguistic cues and pragmatic reasoning – we sought to harness the far-reaching potential of remote methods to explore the generalizability of the phenomenon in question with a larger and more diverse sample. To this end, we recruited 204 children (ages 4.5– 8) from different countries around the world (75.00% from the United States, 20.59% from the United Kingdom, 2.94% from Canada, and less than 1% from Mexico, Australia, and New Zealand). Even within the United States and the United Kingdom, we were able to achieve a high degree of regional diversity (see Figure 6).

Our findings from this first unmoderated remote study largely replicated the original findings of Rhodes et al. (2012) despite a weaker linguistic manipulation than was implemented in the original (see https://osf.io/acrwq/for details). That is, exposure to generic language increased essentialist beliefs, in a manner very similar to the original project (looking at a composite measure of essentialism, children exposed to generic language were 1.6 times as likely to subsequently endorse an essentialist response in both projects). Additionally, the larger sample size of our remote study ($N = 204$ compared to $N = 46$) allowed us to assess which individual components of essentialism drove the overall effect of generic language – a question that was not (and could not) be addressed in the original study. Further, we found consistent effects of generic language on social essentialism when we analyzed our data by geographic region (United States vs. United Kingdom), as well as when we ran the analyses with trials marked as involving parental interferences included or excluded from the sample. Thus, our first use of unmoderated remote research showcased the validity of this platform as a method of data collection for development science, and also illustrated its wide-ranging potential for discovery and innovation within the field.

As another illustration, we are also interested in harnessing the potential of unmoderated remote research as an opportunity to observe parents and children interacting in a home setting. We are currently evaluating this potential via an unmoderated remote study on the role of parents in shaping their children's conceptual development. In the first phase of this project, we are running a multi-session study to conceptually replicate an in-person study examining parent-child conversations about gender (Gelman, Taylor, Nguyen, Leaper, &
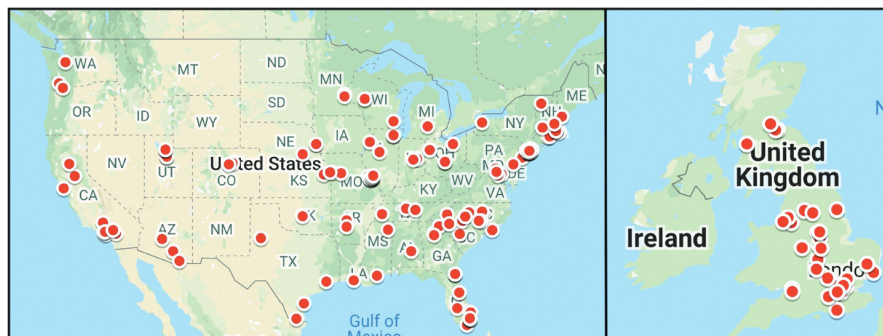

Figure 6. Regional diversity of participants across the United States and United Kingdom from our lab's first remote study (Leshin et al., 2020).

Bigler, 2004). In the first session, parents and children are shown pictures of people engaged in gender-typical and -atypical activities, with a written prompt asking who does each activity (see Figure 2). Parent-child dyads are asked to discuss the pictures as they would do normally. We then code the webcam videos of each family's conversation for various aspects of speech, including how often they use gender labels or generic language (e.g., "Boys play superheroes"), and we also code for various nonverbal behaviors. We then invite families to participate in two additional study sessions, one week apart: first a session measuring children's gender beliefs, which they complete on their own (see Figure 1), followed by a session measuring parents' gender beliefs (see https://osf.io/htn7z/for more about the study). Unmoderated remote research reduces logistical barriers to these types of multi-session studies, enabling us to recruit families with a broad range of gender-related beliefs and experience, and allowing parents and children to discuss gender in their homes, instead of in a lab setting. Therefore, we think this project has the potential to illustrate the power of unmoderated remote research for advancing the field, beyond its capacity to efficiently carry out studies that could have otherwise been done in-person.

Unmoderated remote research and the future of the field

Although we think remote research has great potential, it cannot – and is not meant to – replace in-person research in developmental science. There are many research questions that can only be addressed with studies that involve live, in-person interaction, including those related to social interaction, motor development, spatial reasoning, and so on. There are also some studies that require in-person interaction with an experimenter to help children through the study process in a manner that cannot be recreated online, even with clever animations. Also, there may be no substitute for the powerful role that inter-acting with children directly can play in generating new research questions and hypotheses or in refining new methods. Further, remote research is not an easy or automatic solution for solving problems related to the lack of diversity in the samples included in develop-mental research – we think it has the potential to help address this problem, but we and the field more generally still need to figure out how to best structure our research and institu-tions to ensure the diversity of our samples and generalizability of our results.

These limitations notwithstanding, unmoderated remote research has the potential to advance the field as a tool for facilitating more robust, replicable, and generalizable studies, and as an opportunity for opening new research questions about the role of context in cognitive development. Many labs are making progress on reducing the technological and logistical challenges of remote research in developmental science. We have focused here on an unmoderated approach primarily for verbal children, but others are also pioneering approaches for unmoderated research with infants (Scott et al., 2017; Scott & Schultz, 2017) and methods for remote research that use video-conferencing to preserve interaction with live experimenters (Gweon, Sheskin, Chuey, & Merrick, 2020; Sheskin & Keil, 2018; Sheskin et al., 2020). There is great potential to advance the field with large-scale collaborative approaches to developing these new methodological tools (Sheskin et al., 2020). It is our hope that remote research can help the field address new research questions and advance science in ways that might not otherwise be possible, both when in-person research is restricted and more generally as we move forward as a field.

Acknowledgments

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

References

Adolph, K. E., Gilmore, R. O., & Kennedy, J. L. (2017). Video data and documentation will improve psychological science. *Psychological Science Agenda*. Retrieved from https://www.apa.org/science/about/psa/2017/10/video-data

Adolph, K. E., Kretch, K. S., & LoBue, V. (2014). Fear of heights in infants? *Current Directions in Psychological Science*, *23*(1), 60–66. doi:10.1177/0963721413498895

Adolph, K. E., Vereijken, B., & Shrout, P. E. (2003). What changes in infant walking and why. *Child Development*, *74*(2), 475–497. doi:10.1111/1467-8624.7402011

Borman, K. M. (1981). *Children's interpersonal relationships, playground games and social cognitive skills: Final report*. Washington, DC: National Institute of Education.

Busch, J. T. A., Watson-Jones, R. E., & Legare, C. H. (2018). Cross-cultural variation in the development of folk ecological reasoning. *Evolution and Human Behavior*, *39*(3), 310–319. doi:10.1016/j.evolhumbehav.2018.02.004

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*. doi:10.1038/nrn3475

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: The MIT Press.

Chouinard, B., Scott, K., & Cusack, R. (2019). Using automatic face analysis to score infant behaviour from video collected online. *Infant Behavior and Development*, *54*, 1–12. doi:10.1016/j.infbeh.2018.11.004

Cimpian, A., & Salomon, E. (2014). The inherence heuristic: An intuitive means of making sense of the world, and a potential precursor to psychological essentialism. *Behavioral and Brain Sciences*, *37*(5), 461–480. doi:10.1017/S0140525X13002197

Dahl, A. (2017). Ecological commitments: Why developmental science needs naturalistic methods. *Child Development Perspectives*, *11*, 79–84. doi:10.1111/cdep.12217

Databrary. (2012). *The Databrary Project: A video data library for developmental science*. New York: New York University. Retrieved from http://databrary.org

DeLoache, J. S. (1987). Rapid change in the symbolic functioning of very young children. *Science*, *238*(4833), 1556–1557. doi:10.1126/science.2446392

Engelmann, J., Herrmann, E., & Tomasello, M. (2012). Five-year olds, but not chimpanzees, attempt to manage their reputations. *PloS One*, *7*(10), e48433. doi:10.1371/journal.pone.0048433

Foster-Hanson, E., Cimpian, A., Leshin, R. A., & Rhodes, M. (2020). Asking children to "be helpers" can backfire after setbacks. *Child Development*, *91*(1), 236–248. doi:10.1111/cdev.13147

Gelman, S. A., Taylor, M. G., Nguyen, S. P., Leaper, C., & Bigler, R. S. (2004). Mother-child conversations about gender: Understanding the acquisition of essentialist beliefs. *Monographs of the Society for Research in Child Development*, *69*(1), i–142. JSTOR.

Gilmore, R. O., Kennedy, J. L., & Adolph, K. E. (2018). Practical solutions for sharing data and materials from psychological research. *Advances in Methods and Practices in Psychological Science*, *1*(1), 121–130. doi:10.1177/2515245917746500

Gureckis, T. M., Martin, J., McDonnell, J., Rich, A. S., Markant, D., Coenen, A., … Chan, P. (2016). psiTurk: An open-source framework for conducting replicable behavioral experiments online. *Behavior Research Methods*, *48*(3), 829–842. doi:10.3758/s13428-015-0642-8

Gweon, H., Sheskin, M., Chuey, A., & Merrick, M. (Producers). (2020). *Video-chat studies for online developmental research: Options and best practices* [Webinar]. Retrieved from http://sll.stanford.edu/docs/Webinar_materials_v2.pdf

Hamann, K., Bender, J., & Tomasello, M. (2014). Meritocratic sharing is based on collaboration in 3-year-olds. *Developmental Psychology*, *50*(1), 121–128. doi:10.1037/a0032965

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. doi:10.1017/S0140525X0999152X

Leshin, R., Leslie, S.-J., & Rhodes, M. (2020). *Does it matter how we speak about social kinds? A large, pre-registered, online experimental study of how language shapes the development of essentialist beliefs*. New York University. doi:10.31234/osf.io/nb6ys

Lingeman, J., Freeman, C., & Adolph, K. E. (2014). Datavyu (Version 1.2.2) [Software]. Retrieved from http://datavyu.org

Liu, D., Wellman, H. M., Tardif, T., & Sabbagh, M. A. (2008). Theory of mind development in Chinese children: A meta-analysis of false-belief understanding across cultures and languages. *Developmental Psychology*, *44*(2), 523–531. doi:10.1037/0012-1649.44.2.523

Lourenco, S. F., & Tasimi, A. (2020). No participant left behind: Conducting science during COVID-19. *Trends in Cognitive Sciences*, *24*, 583–584. doi:10.1016/j.tics.2020.05.003

Mandalaywala, T. M. (2020). Does essentialism lead to racial prejudice?: It's not so black and white. In M. Rhodes (Ed.), *Advances in child development and behavior*. Cambridge,MA: Academic Press.

Martin, C. L., & Fabes, R. A. (2001). The stability and consequences of young children's same-sex peer interactions. *Developmental Psychology*, *37*(3), 431–446. doi:10.1037//0012-1649.37.3.431

Medin, D., Waxman, S., Woodring, J., & Washinawatok, K. (2010). Human-centeredness is not a universal feature of young children's reasoning: Culture and experience matter when reasoning about biological entities. *Cognitive Development*, *25*(3), 197–207. doi:10.1016/J.COGDEV.2010.02.001

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, *198*(4312), 75–78. doi:10.1126/science.198.4312.75

Moty, K., & Rhodes, M. (in press). The unintended consequences of the things we say: What generics communicate to children about unmentioned categories. *Psychological Science*. doi:10.31234/osf.io/zkjyr

Nielsen, M., Haun, D., Kärtner, J., & Legare, C. H. (2017). The persistent sampling bias in developmental psychology: A call to action. *Journal of Experimental Child Psychology*, *162*, 31–38. doi:10.1016/j.jecp.2017.04.017

Noyes, A., & Keil, F. C. (2019). Generics designate kinds but not always essences. *Proceedings of the National Academy of Sciences*, *116*(41), 20354–20359. doi:10.1073/pnas.1900105116

Pfungst, O. (1911). *Clever Hans: (The Horse of Mr. Von Osten.) a contribution to experimental animal and human psychology*. New York, NY: Holt, Rinehart and Winston.

Piaget, J. (1954). *The construction of reality in the child* (pp. xiii, 386). Basic Books. doi:10.1037/11168-000

Rhodes, M., Leslie, S.-J., Bianchi, L., & Chalik, L. (2018a). The role of generic language in the early development of social categorization. *Child Development*, *89*(1), 148–155. doi:10.1111/cdev.12714

Rhodes, M., Leslie, S.-J., Saunders, K., Dunham, Y., & Cimpian, A. (2018b). How does social essentialism affect the development of inter-group relations? *Developmental Science*, *21*(1), e12509. doi:10.1111/desc.12509

Rhodes, M., Leslie, S.-J., & Tworek, C. M. (2012). Cultural transmission of social essentialism. *Proceedings of the National Academy of Sciences*, *109*(34), 13526–13531. doi:10.1073/pnas.1208951109

Rizzo, M., Britton, T., Jamieson, K., & Rhodes, M. (2020). *Metrics of racial inequalities*. New York University. Retrieved from https://osf.io/m2evf/

Rosenthal, R. (1966). *Experimenter effects in behavioral research*. New York, NY: Appleton-Century-Crofts.

Ross, N., Medin, D., Coley, J. D., & Atran, S. (2003). Cultural and experiential differences in the development of folkbiological induction. *Cognitive Development*, *18*(1), 25–47. doi:10.1016/S0885-2014(02)00142-9

Rubio-Fernández, P. (2019). Publication standards in infancy research: Three ways to make Violation-of-Expectation studies more reliable. *Infant Behavior and Development*, *54*, 177–188. doi:10.1016/j.infbeh.2018.09.009

Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind: A comparison of Chinese and U.S. preschoolers. *Psychological Science*, *17*(1), 74–81. doi:10.1111/j.1467-9280.2005.01667.x

Scott, K., Chu, J., & Schultz, L. (2017). Lookit (Part 2): Assessing the viability of online developmental research, results from three case studies. *Open Mind*, *1*(1), 15–29. doi:10.1162/OPMI_a_00001

Scott, K., & Schultz, L. (2017). Lookit (Part 1): A new online platform for developmental research. *Open Mind*, *1*(1), 4–14. doi:10.1162/OPMI_a_00002

Sheskin, M., & Keil, F. (2018). *TheChildLab.com a video chat platform for developmental research*. doi:10.31234/osf.io/rn7w5

Sheskin, M., Scott, K., Mills, C. M., Bergelson, E., Bonawitz, E., Spelke, E. S., Fei-Fei, L., Keil, F. C., Gweon, H., Tenenbaum, J. B., Jara-Ettinger, J., Adolph, K. E., Rhodes, M., Frank, M. C., Mehr, S. A., & Schulz, L. (2020). Online developmental science to foster innovation, access, and impact. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2020.06.004

Tamis-LeMonda, C. S., Custode, S., Kuchirko, Y., Escobar, K., & Lo, T. (2019). Routine language: Speech directed to infants during home activities. *Child Development*, *90*(6), 2135–2152. doi:10.1111/cdev.13089

The Many Babies Consortium. (2020). Quantifying sources of variability in infancy research using the infant-directed speech preference. *Advances in Methods and Practices in Psychological Science*, *3*, 24–52. doi:10.1177/2515245919900809

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*, *311*, 1301–1303. doi:10.1126/science.1121448

Wu, Y., Muentener, P., & Schulz, L. E. (2017). One- to four-year-olds connect diverse positive emotional vocalizations to their probable causes. *Proceedings of the National Academy of Sciences*, *114*(45), 11896–11901. doi:10.1073/pnas.1707715114

Xu, K., Nosek, B., & Greenwald, A. (2014). Psychology data from the race implicit association test on the project implicit Demo website. *Journal of Open Psychology Data*, *2*(1), e3. doi:10.5334/jopd.ac