Longform recordings of everyday life: Ethics for best practices

Margaret Cychosz[1] · Rachel Romeo[2] · Melanie Soderstrom[3] · Camila Scaff[4] · Hillary Ganek[5] · Alejandrina Cristia[6] · Marisa Casillas[7] · Kaya de Barbaro[8] · Janet Y. Bang[9] · Adriana Weisleder[10]

Abstract
Recent advances in large-scale data storage and processing offer unprecedented opportunities for behavioral scientists to collect and analyze naturalistic data, including from underrepresented groups. Audio data, particularly real-world audio recordings, are of particular interest to behavioral scientists because they provide high-fidelity access to subtle aspects of daily life and social interactions. However, these methodological advances pose novel risks to research participants and communities. In this article, we outline the benefits and challenges associated with collecting, analyzing, and sharing multi-hour audio recording data. Guided by the principles of autonomy, privacy, beneficence, and justice, we propose a set of ethical guidelines for the use of longform audio recordings in behavioral research. This article is also accompanied by an Open Science Framework Ethics Repository that includes informed consent resources such as frequent participant concerns and sample consent forms.

Keywords Longform recording · Naturalistic · Ethics · Privacy · Confidentiality · Data management

Introduction

Audio recordings have long been a staple of social science research, offering fast, reliable records of human behavior and interaction (Bucher, Fritz, & Quarantelli, 1956; Fasick, 1977; Lee, 2004). Recent years have seen the rise of a particular recording methodology, *real-world longform audio recording*, that captures an audio environment over the course of multiple hours. These recordings (sometimes also known as "long-format" or "[naturalistic] daylong" recordings) are made using a small, lightweight recording device, often in a shirt pocket or small pouch. These devices capture naturalistic human activity for extended periods of time, and may reduce the impact of observer effects on behavior. The recordings provide high-fidelity access to many aspects of research participants' daily life (de Barbaro, 2019; Harari, Müller, Aung, & Rentfrow, 2017, Soderstrom & Wittebolle 2013).

The newfound popularity of this recording method can be partially attributed to technological advances that make it possible to capture, store, and analyze large volumes of audio data. Small, portable devices for audio (and video-) recording are increasingly available and used. Data storage and

Melanie Soderstrom, Camila Scaff, Hillary Ganek, Alejandrina Cristia, Marisa Casillas, Kaya de Barbaro and Janet Y. Bang contributed equally to this work

✉ Margaret Cychosz
mcychosz@berkeley.edu

✉ Adriana Weisleder
adriana.weisleder@northwestern.edu

1 Department of Linguistics, University of California, 1203 Dwinelle Hall, Berkeley, CA 94720, USA

2 Boston Children's Hospital and Massachusetts Institute of Technology, Boston, MA, USA

3 Department of Psychology, University of Manitoba, Winnipeg, MB, Canada

4 Human Ecology Group, Institute of Evolutionary Medicine, University of Zurich, Zürich, Switzerland

5 The Hospital for Sick Children, Toronto, ON, Canada

6 Laboratoire de Sciences Cognitives et de Psycholinguistique, Département d'études cognitives, ENS, EHESS, CNRS, PSL University, Paris, France

7 Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

8 Department of Psychology, The University of Texas at Austin, Austin, TX, USA

9 Department of Psychology, Stanford University, Stanford, CA, USA

10 Department of Communication Sciences and Disorders, Northwestern University, 2240 Campus Dr., Frances Searle Building, Room 3-358, Evanston, IL 60208, USA

processing are also becoming more efficient: automated, open-source processing techniques are beginning to tackle tasks such as speaker identification or speech vocalization counts in hyper-naturalistic settings (Le Franc et al., 2018). In the field of language development, this recording methodology is best exemplified by the advent of the Language ENvironment Analysis system (LENA, Xu, Yapanel, & Gray, 2009), an out-of-the-box audio recording system that includes specialized digital language processing devices (portable audio recorders) and proprietary audio processing software. The ubiquity of mobile phones has also made accessible software solutions that can record audio "snippets" from participants' own devices (Mehl, 2017).

In addition to these technological advances, a number of theoretical considerations in the behavioral sciences have influenced the use of longform audio recordings. First, behavioral scientists have long noted the problems inherent in many data collection techniques (i.e., observer's paradox) and the need for ecologically valid data collected outside formal laboratory environments (Gardner, 2014; Tamis-LeMonda, Kuchirko, Luo, Escobar, & Bornstein, 2017; Mehl, 2017). Second, recognition of extreme sampling bias in behavioral science toward participants from particular populations with unique characteristics (sometimes referred to as "WEIRD" [Western, Educated, Industrialized, Rich, and Democratic]) has led to numerous calls for including data from a wider range of cultural, linguistic, and geographic settings (Henrich, Heine, & Norenzayan, 2010; Nielsen, Haun, Kartner, & Legare, 2017). Third, the replicability crisis in psychological science is increasingly demanding better open-science practices, including greater data sharing and larger sample sizes (Munafò et al., 2017; Stanley, Carter, & Doucouliagos, 2018). Longform recordings deliver on many of these needs: lightweight recording devices permit the collection of naturalistic big data in diverse ecologies, easily collected outside the formal laboratory environment, and audio/video data repositories such as HomeBank (VanDam et al., 2016) and Databrary (Databrary, 2012) facilitate long-term storage and efficient sharing of these recordings with the broader research community.

The same characteristics that make longform recordings so versatile (e.g., unobtrusiveness) also give rise to concerns about privacy and confidentiality. Voice data carries risk of identification,[1] and loss of confidentiality may have more serious consequences when recordings capture real-life behaviors over multiple hours. Thus, researchers using longform audio recordings face challenges related to the ethics of collecting, processing, and archiving these data (Casillas & Cristia, 2019). In addition, as the scope of populations under study broadens, researchers need to reflect on how to

guarantee that the dignity of all participants and groups is respected. This is true of all research, but is especially important in the context of methodologies like longform audio recordings that have the potential to capture and reveal information that individuals or groups would not have willingly revealed – especially for members of marginalized groups.

Recent guidelines have been proposed for ethical research practices involving video data in real-world environments (e.g., using wearable cameras; Derry et al., 2010; Kelly et al., 2013). These guidelines extend ethical frameworks for human subjects research by considering novel issues related to the passive capture of large volumes of video data in authentic environments. While there are many similarities in the issues that arise from real-world video and audio research, there are also important differences by modality (visual vs. auditory), the technology available for capturing and analyzing video and audio data, and participants' perceptions of privacy when wearing video or audio recorders (Alharbi et al., 2018). Here, building on fundamental principles of human subjects research and previous guidelines developed for video research (Derry et al., 2010; Kelly et al., 2013), we discuss ethical considerations for studies that use longform audio recordings and propose a set of guidelines for ethical research using this methodology.

We are an international group of researchers with significant experience using longform audio recordings. Although not legal scholars or bioethics experts, we have collectively been involved in more than 15 successful ethics board applications involving research with longform recordings. We are often contacted for guidance by other researchers wishing to use this methodology and believe that this work may serve as a useful guiding tool. Our primary target audience for this article is researchers who may be contemplating the use of longform audio recordings in their research or are just starting out. The best time to implement these practices is right from the beginning of a research program. However, researchers with more experience may also benefit from the framework that we outline here. Indeed, the development of this paper highlighted the need for change within some of our own practices, and we envision this paper as the beginning of a broader discussion of best practices in the research community rather than the final word. To that end, this article is accompanied by an Open Science Framework (OSF) Repository that contains sample consent forms and scripts, participant FAQs, and links to relevant resources discussed in this article (Cychosz et al., 2020).

Because the ethics underlying this research are complex and best practices are situational, we do not simply give a fixed set of guidelines. Rather we begin, in section 2, by briefly examining the broader internationally recognized principles of human subjects research, with a focus on application to longform audio recordings. In section 3, we discuss the benefits and risks of this research to participants and society at large, as well as to researchers and the scientific community. Drawing on these considerations, in section 4, we propose a

---

[1] Note that ethics review boards and granting agencies vary on whether they consider voice data to be de-identifiable.

set of guidelines for ethical research practices using longform audio recordings (summarized in Tables 1 and 2) and suggestions when obtaining informed consent (Table 3).

Guiding principles

Universal principles for ethical human subjects research

The Declaration of Helsinki, published originally in 1964 and subsequently amended (World Medical Association, 2013), outlines the key principles that have been used to set the standards for most research involving human subjects around the world today. These include respect for autonomy, respect for privacy and confidentiality, beneficence, justice, and protection of vulnerable populations. Here, we briefly summarize these principles and their application to longform audio recordings.

One of the principles of human subjects research centers on respecting the *autonomy* of research participants and allowing participants' self-determination. This is primarily achieved through the process of informed consent, in which participants are informed of the goals of the research and the risks and benefits that can result from participation before agreeing to participate. Longform audio recordings can present unique challenges for respecting autonomy and ensuring informed consent because of the many possible types and uses of the data (see Fig. 1), and the possibility of recording third parties who have not had direct communication with researchers.

Two other key principles focus on participants' right to control access to their information (*privacy*) and protect their identifiable information (*confidentiality*). This is perhaps one of the greatest concerns with the use of longform audio data, given the potential for recording sensitive information, as well as unlikely but possible mandates to break confidentiality for various legal reasons, which must be considered in ethical research plans.

The principle of *beneficence* refers to the philosophy to "do no harm". In other words, researchers must maximize the benefits of the research project while simultaneously minimizing risk to research subjects at all stages of the project, including data collection, analysis, storage, sharing, and dissemination.

Finally, the principle of *justice* demands that the risks and benefits of scientific inquiry be fairly distributed, ideally involving the same type of participants who would benefit from the results of the research. With respect to longform audio recordings, this principle entails that the knowledge, practices, or interventions that result from the research should find their way back to the relevant groups.

Further considerations for vulnerable populations

Additional considerations are necessary for research involving vulnerable groups, such as minors, pregnant women, prisoners, employees, people with disabilities, ethnic and racial minorities, and those who are economically disadvantaged (Office for Human Research Protections, 2019). Traditionally, extra protections have been extended to these groups to prevent exploitation of individuals who have a dependent status or otherwise impaired capacity to give informed consent. However, identifying certain groups as vulnerable can also lead to stereotyping and excessive paternalism, and result in exclusion of some groups from research (Rogers and Lange, 2013). Ethical considerations for vulnerable groups must balance two opposing concerns: on the one hand, some groups may be overrepresented in research because of a lack of understanding or power to refuse participation; on the other hand, overprotection of some groups may result in these groups being excluded from research and from the benefits that could result from research participation (Nickel, 2006).

For example, individuals with low literacy and low technological literacy may have greater difficulty understanding the possible uses of longform recordings from an informed consent form, and thus may unintentionally agree to exploitative terms. To protect
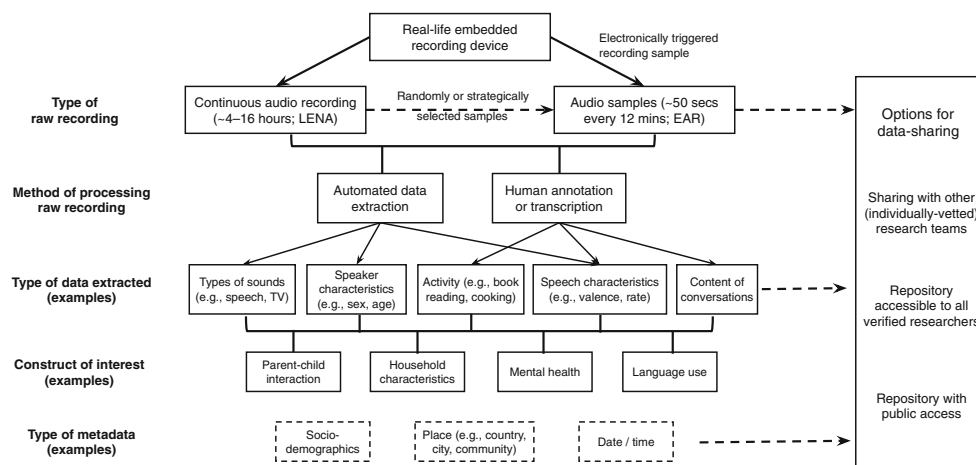

Fig. 1 Research pipeline from collecting longform audio recordings to extracting data that can be used to measure the construct(s) of interest. Decisions made at each point in this pipeline require consideration of the risks and benefits involved. Dashed lines indicate optional steps

against exploitation, researchers could avoid conducting this research in communities with high rates of illiteracy. Yet excluding these communities may mean that any knowledge stemming from research on longform recordings – including the development of speech recognition tools that could improve access to information, or standardized language metrics that may be used to gauge developmental progress – will not represent those groups and may not benefit those communities, exacerbating existing disadvantages. Researchers should instead find alternative ways to seek informed consent from participants with low literacy and those who have limited familiarity with recording technologies, avoiding both exploitation and exclusion.

In the sections below, we make efforts to consider the impact of various research practices on groups that have historically been underrepresented in behavioral research, including ethnic and racial minorities, individuals living in non-western societies, and/or individuals with low levels of education or income. Although a full discussion of ethical issues in research with vulnerable populations is beyond the scope of this paper (for useful discussions, see Hill, 1995 and Talbert, 2019), our goal is to center considerations for vulnerable populations within a discussion of ethical research practices for longform recordings.

Risks and benefits in research with longform audio recordings

Figure 1 illustrates the progression from collecting longform audio recordings to extracting data that can be used to measure the construct(s) of interest. Decisions made at each point in this pipeline – from how to collect and process the recording, to the type of data extracted and constructs analyzed – require consideration of the risks and benefits involved. For example, analysis of different aspects of participants' behaviors and interactions could be performed based on short audio clips (e.g., 50 seconds) sampled at regular intervals. This sampling could be done by the device during the recording process, as with the Electronically Activated Recorder (EAR; Mehl & Pennebaker, 2003), or through an algorithm applied to a longer recording after it has been collected. Using short snippets of audio could reduce the risk of capturing personal or sensitive information, and thus may help protect participants' confidentiality. However, this method may not be sufficient when the goal of the research is to capture the dynamics of interactions or analyze how behaviors change over multiple timescales.

Decisions about how to process the recordings to extract usable data also involve weighing risks and benefits. For example, LENA's proprietary software produces annotations about the types of sounds (e.g., noise, electronic media, speech) in a 16-hour recording, and provides estimated counts of adult words and child vocalizations. Researchers can choose to use the automated counts and throw away the raw recordings, thus minimizing the risk to participants from loss of confidentiality.

However, deleting the raw recordings eliminates the possibility of checking that the automatically generated data accurately capture the constructs of interest (e.g., Weisleder & Fernald, 2013) and significantly limits the potential for future analyses.

Similarly complex issues arise when deciding what type(s) of data will be extracted from the recordings, the constructs that will be analyzed, and the type(s) of data that will be shared with other researchers for purposes of verification or secondary analyses. In the sections below, we outline both benefits and risks from research with longform recordings with the goal of helping researchers approach these decisions.

Benefits afforded by real-world longform recordings

We anticipate benefits from this type of research at the level of individual participants and communities, individual researchers, and the scientific community (Table 1).

Benefits to participants and communities

Analyses of large-scale databases of audio recordings may permit a deeper understanding of what is effective in promoting better health and mental health outcomes across populations, with ensuing benefits for participants and communities. For example, in the area of language development, longform audio recordings have the potential to revolutionize public health approaches to promoting language and literacy in young children by providing quantitative measures of children's language environments that can be used both as feedback to caregivers and as a measure of the efficacy of various interventions (Zhang et al., 2015; Ferjan Ramírez, Lytle, Fish, & Kuhl, 2019; Leung, Hernandez, & Suskind, 2019). Parents in the United States report that receiving a variety of feedback from ambulatory recordings – from parenting advice to deviant behaviors – would increase their reported willingness to participate in mobile-sensing studies (Levin, Egger, Johnson, & de Barbaro, 2019), suggesting this type of feedback is valued among U.S. parents. Though these approaches are not without criticism (Avineri et al., 2015), they suggest that longform recordings could have direct benefits for some participants. Moreover, studies of interventions can be important in shaping public policy, thus having broad impacts for communities and society at large (Greenwood et al., 2017).

Collecting naturalistic data in diverse communities is critical for establishing the normative standards and practices used to inspire and measure interventions and the recommendations derived from them. More inclusive research sampling increases the potential for generalization to the broader population, and may also reveal limits in the applicability of research findings to different communities. For example, while research with some populations in the United States has found a relation between children's exposure to child-directed speech and their language development, it is unclear whether the same

Table 1 Sample benefits and risks of longform audio recordings by population

	Benefit	Risk
Individual participants	• More life experiences represented in scientific research • Quantitative reports (e.g., daily language counts) could be made available to participants and clinicians	• Loss of privacy and confidentiality as participants record their personal information, such as names and financial information • Possible legal consequences if illicit activity such as drug use or domestic abuse is recorded, discovered, and reported
Society and communities at large	• Studies of interventions carried out using the technology may lead to advances in public policy/health recommendations • Comprehensive data sets document cultural and linguistic traditions • Development of normative standards based on more representative samples	• Generalizations of potentially unflattering findings • Potential to reveal aspects of family/community life that conflict with how communities would want to be portrayed • Propagation of bias if algorithms are developed based on nonrepresentative samples
Researcher/research team	• Efficient collection of large quantities of data • Ability to reuse data for additional research projects	• Snapshots of private moments or illegal activity may lead to discomfort or researcher ethical conflict of interest • Potential for equipment damage as recordings are made outside the lab
Scientific community	• Larger, more comprehensive data sets, leading to greater ecological validity and generalizability of findings • Documentation of naturalistic behaviors from more diverse communities, leading to robust theories • Processing algorithms that permit comparison across data sets and laboratories • Greater opportunities for data reuse among researchers with varied interests from different fields	• Processing algorithms that permit comparison across data sets and laboratories may not capture relevant behaviors within individual communities • Overreliance on audio could lead to theories that emphasize auditory and verbal information over other modalities • Exploring the vast number of measures that can be derived could lead to spurious findings

relationship exists, or would be useful, in countries or cultures with different ideologies about caregiving and cognitive development (Casillas, Brown, & Levinson, 2019; Cristia, Dupoux, Gurven, & Stieglitz, 2017). Research with longform audio recordings enables us to test the generalizability of such findings prior to making global recommendations.

Audio recordings have additional community benefits outside of healthcare policy and practice. For indigenous communities and/or communities speaking an under-documented or endangered language, these recording technologies may be an efficient means for creating a large body of natural language data, caregiving practices, and other verbally expressed traditional practices. Longform recordings could serve as a snapshot of linguistic and cultural practices that may be changing in response to globalization. In the case of Ōlelo Hawai i, the precolonial language of Hawai'i, audio recordings helped document and preserve the language when its use was declining; recordings were used to create teaching and learning materials to revitalize the language (Brenzinger & Heinrich, 2013). But even if there is no current risk of cultural loss, such ecological recordings may better enable others to view such practices in their natural context. Through the recognition and preservation of diversity in human behavioral practices, we have an increased ability to address the significant societal and humanitarian issues that many people from underrepresented groups disproportionately face. The improvement of the annotation tools accompanying these technologies also has the potential to help maintain and document linguistic and

cultural diversity, particularly the richness of conversational language that is not always captured in writing.

Finally, incorporating longform recording methodologies into participatory research models could also lead to more direct benefits for the communities being studied (Gehlert & Mozersky, 2018). Community-based participatory research seeks to center the needs and interests of the community when setting the research agenda (Luo et al., 2019). Thus, community members help generate research questions, recruit other participants, and analyze data. Community member involvement in research increases trust and transparency; communicating the intentions and risks of a research study is much easier when a community member who has participated in data collection, annotation, and/or analysis can directly share their expertise with other participants. This process gives the community a voice in how the research is done, and makes the topics under study more visible to the participating families.

An example of the community-based approach would be if researchers convened focus groups from the community and used their discussions to inform research questions about communication behaviors that are considered to be valuable for that particular cultural community. Researchers can also work closely with a few individuals who are interested in learning more about the research topic. As these individuals gain expertise in the technical and theoretical aspects of the research, they may provide influential insights that change the course of the work for the better (e.g., noticing a pattern in behavioral data that is not immediately visible to the researcher), even making contributions as co-authors. Thus, participatory

research can help mitigate power differentials between the community and the researchers and ensure that community interests are represented in the products of scientific discovery.

Benefits to researchers and the scientific community

For individual researchers and research teams, the benefits of longform recording technologies are clear: researchers can collect more data from natural environments quickly and more consistently than previously possible. Even when these recordings are not the central data for a research project, they can be used to complement other modes of investigation, from in-lab experiments to qualitative microanalyses. And, once collected, they can be flexibly reused for a variety of future research questions.

Researchers also stand to benefit from sharing their data with others. For individual researchers, the benefits of sharing longform recording data include: membership in like-minded research communities who may be facing (and solving) the same kinds of problems; the ability for data to be reanalyzed and enriched by others, thereby facilitating future analyses; and a framework for data storage and access during researcher transitions (e.g., moving institutions) or after researcher retirement. Furthermore, collaboration in research groups does not require that members openly share all of their longform recordings. For example, a recent multi-lab collaboration involved the large-scale annotation of small clips (~400 milliseconds extracted from day-long recordings) by citizen scientists (Cychosz et al., 2019). Through the collaboration, large amounts of data from diverse groups were included in the analysis. But the annotation methodology also protected participant privacy and confidentiality since the audio clips were so short. This is one example of collaborative data sharing that protects participants and does not involve sharing entire recordings.

For the larger scientific community, research initiatives will be facilitated by the collection and analysis of larger samples that originate from diverse populations. These developments may pave the way for changes to theories of human development and behavior. For example, the research community can use longform recordings to test core claims on newer, broader data sets and build upon theories based on behaviors not previously systematically observed because of limited data. In this context, shared data again multiplies the benefits because it means that findings will be open to replication, improved analyses, and added information, increasing scientific accuracy and efficiency (researchers derive more findings from data sets at a faster pace). Shared data also spurs on the continued development of tools for data collection, annotation, and analysis through a larger community framework. More output in this domain will also motivate new innovations, such as improvement of tools for visual information, geographical data, or psychophysical states that accompany the audio recording, moving toward a more holistic view of the participant's experience in the world.

Risks involved in research with longform audio recordings

Longform naturalistic recordings can also lead to risks to individual participants and communities, as well as the researchers conducting the work and the scientific community (Table 1).

Risks to participants and communities

The risks to participants in research using longform audio recordings stem primarily from the potential loss of privacy and confidentiality, and the possibility that this loss of confidentiality brings harm to individual participants and/or the communities they belong to. This risk is the product of three factors: (a) the likelihood that individuals (or groups) can be identified from the recordings, (b) the degree of sensitive, compromising, or unflattering content in the recordings, and (c) the possibility that identifiable information is disclosed to persons who know the participant or who can have an influence on participants' or communities' lives (e.g., law enforcement, government policymakers). If any of these factors is minimized to zero (or close to zero), the potential harm to participants is negligible. For example, research on highly sensitive topics such as drug use or crime often uses anonymous surveys to ensure that compromising information cannot be linked back to individual participants, thus minimizing any potential harm from sharing this information. At the other end, language acquisition researchers have created shared databases of audio and video corpora in which individuals are identifiable (e.g., CHILDES, Databrary), but which typically showcase interactions that are relatively constrained and considered to be harmless (e.g., video recordings of parent-child play), and thus carry minimal risk to participants.

A recent survey of parents' willingness to collect and share mobile-sensing data of children's daily experiences showed that a representative sample of U.S. parents queried in 2018 ($N = 210$) were less comfortable collecting audio (and video) data than they were collecting motion and heart rate data when provided with a brief study description. In particular, 41% of parents were at least "somewhat willing" to share audio recordings, and only 15% of parents reported being "extremely willing" to collect audio data (Levin et al., 2019). These data suggest that participants in the United States do perceive a relatively high risk of loss of privacy from collecting and sharing longform audio recordings, and that audio data is perceived to be more sensitive than other types of ambulatory data

Several factors can affect the risk of identifiability in longform recordings. Although voice data is, strictly speaking, an identifier (Department of Health and Human Services, 2015), in practice, voice alone is only identifying if the speaker is already known to individuals listening to the recording or if their voice can be found in public repositories (e.g., YouTube, SoundCloud) from which the person can be uniquely

identified. This means that for most research participants, the likelihood of identification through voice data alone is extremely low. However, the anonymity of voice recordings cannot be completely guaranteed. First, there is a possibility that the risk of identification will change in the future. For example, it is not impossible that a database of voice recordings (e.g., collected via commercially available smart speakers) could be paired with identifying information (participant names) which could be used to "fingerprint" individuals. While we are unaware whether such a database exists at a truly large scale, using voice to uniquely identify a speaker appears technically feasible. Second, even if a given participant is not known by the research team collecting the data, they could ostensibly be known to researchers with access to the data in the future. Thus, the risk of identifiability grows as the voice data is shared, and is difficult to meaningfully measure. In addition, it is possible that information contained in the recordings themselves could identify the participant or others in the recording (e.g., use of names or other personal information).

Because longform recordings are intentionally unobtrusive, it is also likely that participants will say or do things that they would not wish others to hear, such as disclosing financial information or embarrassing content. Thus, in general, there is a non-negligible risk of harm to participants if there were a breach of confidentiality, which can occur if researchers do not take sufficient precautions to safeguard the data through implementation of appropriate staff training and secure data storage and transfer. This risk could increase as data are accessed from and potentially stored in multiple sites. In some cases, longform recordings could also include evidence of illegal activity, such as drug use, domestic violence, or child abuse. Many researchers are mandated reporters and would be required to report these instances to authorities (see section 4.2), constituting a break in confidentiality. This possibility needs to be disclosed as a risk to participants in the informed consent process, as we discuss in detail below.

Even when participants' individual identities are kept confidential, disclosing information about group characteristics (e.g., race, ethnicity, nationality, gender) could result in risks to the communities represented. Although these risks are not unique to longform recordings, these recordings have the potential to reveal aspects of family and community life that are frequently kept private and that could be misinterpreted by individuals from outside the communities studied. Generalizations drawn from these findings may conflict with how communities view themselves and how they would want to be publicly represented. This risk is higher for marginalized groups, who are more likely to be stereotyped and stigmatized.

Finally, as with any methodology, a potential risk to participants and communities stems from how the data are interpreted. Although longform recordings represent an improvement over some recording methodologies because of the wider scope of behaviors they capture, this method is not

without potential for bias. In particular, if normative standards are developed based on recordings from nonrepresentative samples, this may result in a distorted picture of what is considered "typical" human behavior and development. Researchers should be cautious when interpreting data from longform recordings to avoid normalizing behaviors and experiences that are characteristic of only a small cross-section of the population (often one with more power and privilege) and thus stigmatizing the behaviors and practices of underrepresented groups. In addition, because individuals from marginalized groups are underrepresented in the scientific community, there is a higher risk that the behaviors and practices of marginalized groups will be misinterpreted, and a lower likelihood that the community's wishes will be represented in the formulation of research questions and in the development of interventions and policies that could affect those communities. As discussed above, a participatory research framework can help reduce some of these risks.

Risks to researchers and the scientific community

As discussed above, longform recordings may contain violence and other emotionally charged events. Researchers listening to the recordings for annotation or transcription could come into contact with such segments, which could cause them distress. Decisions about reporting apparent illegal activity or perceived/possible child abuse can also cause researchers to experience ethical conflicts. Researchers who will be listening to longform recordings should receive appropriate training in order to protect both the participants and the researchers themselves (see section 4.3).

Risks to the scientific community include an overreliance on data emerging from longform audio recordings. When used properly, these data sets will be large and supposedly comprehensive and ecologically valid, making it tempting for some to believe that they capture individuals' behavior at its most natural. However, everyone has days or even weeks that are atypical, and it is unclear how long a recording would have to be to allow a truly representative view of an individual's environment (Anderson & Fausey, 2019). There are also significant limitations to the behaviors and interactions that can be captured by audio recordings. For example, reliance on this methodology could lead to the development of theories that overemphasize the role of auditory and verbal information over other modalities.

Moreover, recordings per se do not constitute interpretable data: Coding, transcription, or processing algorithms need to be applied to convert the data into interpretable measures. The field of machine learning is teeming with examples of algorithm bias, which occurs when algorithms are developed with biased data sets (Garcia, 2016). If audio processing algorithms are developed based on the narrow samples that are typical of psychological research (Arnett, 2008), these algorithms may introduce bias when used to analyze data from a broader range

of populations. Researchers should be cautious when applying existing algorithms to recordings that include languages or contexts different than the ones for which the algorithms were developed (Canault, Le Normand, Foudil, Loundon, & Thai-Van, 2016; Ganek, Smyth, Nixon, & Eriks-Brophy, 2018) and, in the longer term, should seek to develop algorithms based on more diverse samples (Cristia, Ganesh, Casillas, & Ganapathy, 2018; Räsänen et al., 2019). A final risk to the scientific community stems from the size and multivariate nature of the measures that can be derived from such data sets. Unless used carefully, carrying out statistical analyses on this vast number of measures could lead to spurious findings.

Considerations and best practices in research with longform audio recordings

Based on the ethical issues discussed above, and building on previous ethical frameworks (Derry et al., 2010; Kelly et al., 2013), we propose the following set of guidelines for ethical research practices using longform audio recordings (Tables 2 and 3). These guidelines include recommendations related to protecting privacy and confidentiality, dealing with sensitive information in recordings, appropriate training for research assistants, sharing and reusing data, and obtaining informed consent from participants and third parties.

Protecting privacy and confidentiality

Protecting the privacy and confidentiality of participants (and third parties) can rely on several strategies that can be implemented before, during, and after the recording process (Langheinrich, 2001). A simple and useful provision is to instruct participants that they can pause or remove the recorder at any time, thus granting them more autonomy over their recordings. Another low-tech option is to give participants a logbook where they can mark sections of the recording or times of day that are to be deleted or simply not listened to by the researcher.[2] Hypothetically, in settings where participants have access to smartphone technology, a QR code could direct participants to a website where they could signal that they would like the previous five minutes of the recording deleted. While currently not an option with LENA, custom-built devices have been developed that directly allow

participants to delete recorded material (Kay et al., 2012; Kiukkonen, Dousse, & Gatica-Perez, 2010).

There are also a number of technological solutions that can increase privacy for research participants for some kinds of research questions. The basis of most of these techniques is that the collection and storage of filtered or low-resolution data will suffice for many analyses. In the domain of location data, for example, collection and storage of neighborhood- or block-level information rather than exact GPS coordinates can provide the necessary information for researchers while also obscuring identifying information. In the audio domain, a number of such techniques have been developed, including feature extraction techniques, subsampling, and distortion of the speech signal to reduce intelligibility while preserving key acoustic features (Kumar, Nguyen, Zeng, Liu, & Zhang, 2015; Chen, Adcock, & Krishnagiri, 2008; Larson, Lee, Liu, Rosenfeld, & Patel, 2011). Note, however, that mention of these privacy-preserving techniques appears to make only a small difference in participants' willingness to collect and share audio data (Levin et al., 2019), suggesting the need to further explain to participants how these techniques can protect their privacy and confidentiality.

Algorithms that automatically extract relevant features from audio data can enhance participant privacy by extracting and storing features of audio rather than raw audio data itself in such a way that conversations or recognizable speaker characteristics would no longer be available (Wyatt, Choudhury, & Bilmes, 2007; Lane et al., 2011). These algorithms could be run on stored audio data or directly on the recording device or data server (as in Wang et al., 2014a, b). In the former case, raw audio files can be deleted; in the latter, raw audio samples may never need to be stored, as data may be processed in real time on the device or remotely on a cloud computing service.

These techniques could greatly enhance participant privacy by eliminating the storage of raw audio recordings. However, some limitations exist. First, when participants pause/remove a recorder, it becomes easy to forget to start/wear the recorder again. Also, getting rid of raw audio restricts the range of questions that can be addressed during audio processing (e.g., acoustic measures of speech, but not most other linguistic analyses). Some of the appeal of audio data is precisely the fact that it is a high-fidelity signal that can provide researchers with endless research possibilities which may not have been anticipated at the start of the study.

Next, while showing proof of concept, most audio algorithms relevant for behavioral scientists have been developed with small samples and would require additional validation efforts (see review by de Barbaro, 2019). Exceptions are the speech recognition algorithms developed by LENA and a few simple algorithms characterizing human voice and presence of conversations (Wyatt et al., 2007; Lane et al., 2011). Thus, much work remains to be done before additional variables could be automatically extracted.

---

[2] Note that the option to delete sections of the recording may not always be technologically possible. For example, the LENA system provides a UPL file, a storage form from which both quantitative and audio data may be generated and a wav file may be extracted. Consequently, while portions of the wav audio file may be deleted, researchers maintain the information about the original recording unless the full UPL file is likewise deleted. To our knowledge, there is currently no means to partially delete information in the UPL files.

Table 2 Ethical guidelines for longform audio recordings in behavioral research

Protection of privacy and confidentiality	*Increasing participant autonomy over recordings*
	- instruct participants on how to remove the device and/or pause the recording
	- provide a mechanism for participants to request deletion of sensitive information from the recordings
	Researcher's role in protecting participants' privacy
	- configure devices so that recordings can be retrieved only by researchers; third parties cannot access data if device gets lost
	- use filtered or low-resolution data when it will suffice for desired analyses (e.g., feature extraction techniques, subsampling)
	- vet recordings for personal information prior to public sharing; stricter vetting procedures may be necessary for third parties
	- provide appropriate privacy and confidentiality training for researchers who will have access to the audio data
Illegal activities and safety concerns in recordings	- ascertain whether the researcher(s) are mandated reporters as defined by an employment contract/applicable laws
	- include information about mandated reporting and qualifying events in the consent form (see Table 3)
	- determine what qualifies as a reportable event, in consideration of applicable laws, time since data collection, and the purported risk to participants or others
	- specify how researchers are to identify reportable events (e.g., are only incidental findings considered or will researchers conduct oversight on all recordings)
	- be aware of the procedure for reporting safety concerns/illegal activity (e.g., identifying the correct person to contact and delineating how much information is required to be disclosed)
	- if reporting is deemed necessary, determine whether to notify participants about the breaking of confidentiality
	- provide appropriate training for research staff, including cultural training that addresses potential for misinterpretation
Data sharing and storage	- make explicit plans for long- and short-term data management
	- do not make promises that you will be unable or unwilling to keep, such as destroying data after a number of years
	- consider different dissemination levels based on length and content of audio snippets from recordings
	- adjust template consent forms to include data sharing and storage options that conform to the needs of your research (see Table 3); follow participants' wishes for how their data should be shared and reused
	- store data according to national data protection regulations
	- be prepared to explain your data storage system in a comprehensible manner to ensure participant understanding
	- choose an appropriate data repository venue (see section 4.4.1)
	- include a meta-data statement that explains the conditions for reuse
Third parties	*Researcher responsibilities prior to and after data collection*
	- be familiar with the third-party consent laws of the country/region in which the research is taking place
	- provide participants with an information card (or short verbal explanation) to give to third parties that explains the device, study goals, and privacy options (e.g., pausing the recording)
	- do not share or publish recordings of third parties without their consent
	Participant responsibilities while recording
	- if recording at home, participants should seek verbal permission from those likely to appear in the recording
	- if socializing with friends and acquaintances, participants should alert them to the recording (can use information card)
	- in public spaces (e.g., market, park), participants can wear a salient marker (e.g., distinctive clothing, a "warning triangle" on the device) that alerts individuals to the recording
Considerations for vulnerable populations	- address potential vulnerabilities to exploitation or exclusion
	o low literacy or language barrier
	o power differentials
	o allure of material incentives in low-resource contexts
	- consider the need for community in addition to individual consent
	- consider risk of increasing stigma or perpetuating stereotypes

Table 3 Information that should be included in the informed consent process for longform audio recordings in behavioral research

General description	- how the recording will be used to address study goals (and possibility of additional study goals; see data reuse below)
	- length and scope of recording (e.g., continuous vs. snippets, recording radius)
	- examples of what could be recorded (e.g., environmental sounds, background conversations)
	- how recordings will be processed (e.g., automated software, researchers listening to short snippets/entire recording)
	- type of data to be extracted (e.g., speech quantity, emotional markers, conversational content)
	- whether the recording will be stored or discarded (and whether participants can opt out of this)
	- other data that will be linked to recordings (e.g., participant demographics, geographic location, date/time)
Risks to privacy and confidentiality	- reminder that device could record sensitive or unflattering interactions (e.g., yelling, using the bathroom)
	- certain activities (e.g., child maltreatment) may not be protected by confidentiality and could be reported to authorities
	- identify others who may access the recordings under certain conditions, e.g. university research oversight, technical staff
	- complete confidentiality cannot be guaranteed (e.g. risk of a data breach, reporting mandates)
Safeguards to protect privacy and confidentiality	- let participants know they are able to remove the device or pause recording whenever they wish
	- detail options for participants to vet recording after it is made (e.g., will there be an opportunity to delete portions of the recording?)
	- describe any options for increasing privacy (e.g., on-device feature extraction)
	- provide assurance that only trained researchers will have access to the recording unless additional permissions are obtained
	- no recordings will be shared in publications/presentations without explicit consent
Data sharing and reuse	- describe any plans for data sharing and whether participants can opt out of them
	- provide tiered options allowing sharing of audio data (e.g., data can be shared (1) with select research teams, (2) on a protected repository available to verified researchers, (3) on a repository with unvetted public access)
	- provide tiered options allowing use of data for additional research questions
Strategies to facilitate and ensure comprehension	- provide examples of audio clips that demonstrate what recordings sound like and what they capture
	- use consent comprehension quizzes to ensure participants understand these details

Furthermore, while simple algorithms have been previously implemented on mobile phones, processing constraints mean that it is not feasible to run more sophisticated algorithms in real time. Additionally, power requirements and built-in systems which prioritize microphone inputs mean that cell phones cannot record and transmit audio data continuously (only in snippets). However, standalone devices similar to the LENA could ostensibly continuously record and process incoming audio data, and there are already some efforts to develop such systems (e.g., Starling, https://getstarling.versame.com/starling/). Additionally, research optimizing hardware and algorithms for on-device or "edge" computing is an active research area (Chen et al., 2017). For cases where this is not yet possible, software-based solutions which encrypt audio data on the device before automatically processing it off-line are also an option.

Another alternative to partially protect participant privacy and confidentiality during the recording process is to only analyze shorts samples from the raw longform audio data, as with the electronically activated recorder (EAR; Mehl & Pennebaker, 2003; Mehl, 2017). By limiting the type and quantity of data that is collected, this solution retains high-fidelity audio data while preserving more of participants'

privacy. However, researchers should be sure not to guarantee security when explaining subsampling processes to participants: shorter clips may appear less invasive, but may still also occasionally contain sensitive information. For example, a 30-second snippet is enough to capture unflattering language or a threat of corporal punishment. Moreover, depending on the sampling technique (e.g., random sampling), this method could simply mean that individuals listening to the clips hear less context. This absence could, among other things, lead to misinterpretation. Additionally, for some questions, subsampling may not be sufficient to capture the dynamics of behaviors that extend over periods longer than a few minutes. One future possibility for intelligently limiting identifiable data is combining algorithms that automatically detect the events or behaviors of interest and then launch recordings for short windows.

After data collection, researchers can take several steps to remove/mask sensitive and identifying information from longform audio recordings. A first step that many researchers take is to vet the recordings, especially prior to sharing (semi-)publicly. A guide for vetting longform recordings can be found among the HomeBank Project's Resources (https://homebank.talkbank.org/; VanDam, Warlaumont,

MacWhinney, Soderstrom, & Bergelson, 2018). This cursory listen to a recording could involve flagging inappropriate language or elements of the recording that the participant would likely not want to be publicly released or, in the case of a child, might find embarrassing at a later date (e.g., potty training). In our collective experience, these events are relatively infrequent in most study contexts. However, most of the authors have come across information that could be embarrassing or damaging to the participants, suggesting the risk of encountering these events is not negligible and supporting the need for a vetting step prior to sharing the recordings. Determining what might be considered inappropriate or embarrassing is challenging to implement systematically, subject to interpretation, and highly culturally defined. Another option is for participants themselves to review portions of the audio recording. Participants could then be given the option to delete some or all of the audio recording upon completion. However, such a vetting step is impractical in the context of extremely large quantities of audio or when families are hard to reach. Moreover, both types of vetting may be beyond the means of labs with limited resources.

A final option that researchers can consider is the level of sharing. Many long-term, web-based data storage systems offer various sharing levels (e.g., fully public, researchers only, individual permission only). Under these circumstances, the extent to which vetting is needed will depend on criteria such as the nature of the consent obtained from the participants, the level of concern over potential harm to individuals or groups, and the ethical training of those getting access to the recordings. Researchers may additionally consider having a stricter vetting procedure for third parties who are incidentally recorded. An example consent form that includes options for different levels of data sharing is included in the Ethics Repository associated with this article (Cychosz et al., 2020).

Since tools for automatically extracting speech information from raw audio are still quite limited, it would be desirable to have more manually annotated data which is easily shared so that speech technology and machine learning experts can improve automatic annotation algorithms. Notice that any of the above recommendations to remove or modify the audio, as well as limit access, will result in less data being available for algorithm (re)training. Thus, short- and long-term techniques for preserving participant privacy may sometimes tug in opposite directions.

Illegal activities and safety concerns in recordings

Because longform audio recordings are meant to sample participants' real-world experiences, it is possible that recordings may contain evidence of illicit activity or material concerning the safety of participants or others (e.g., illegal drug use; physical, sexual, or psychological abuse). Whether researchers should report this information to relevant agencies raises many ethical questions. From a legal perspective, this differs depending on the laws of the researcher's geographical jurisdiction and rules set forth by individual ethics boards that govern the intended research.

In many countries, select professionals are considered *mandatory reporters* who are legally required to report suspected abuse or neglect of children and/or other vulnerable populations, such as dependent adults and the elderly (Dubowitz, Hein, & Tummala, 2018). Mandated reporting professions typically include child care providers, teachers and school personnel, physical and mental health care providers, police and law enforcement, and social workers (DHS, 2016). Although researchers, including graduate students, are not always *explicitly* named as mandated reporters – in the United States, this is more common at public universities – researchers are nevertheless often university teachers. Others may also have a dual role as a clinician. Consequently, courts may find that university researchers fall under mandatory reporter status (Sieber, 1994). Furthermore, some countries and U.S. states require reporting by all professionals and/or civilians, and many jurisdictions allow for voluntary reporting by concerned persons who are not under mandate. Failure of a mandated reporter to report suspected abuse may result in criminal penalties. Even personnel such as research assistants that do not have mandatory reporter status may have an obligation to report instances of abuse to the lab's principal investigator.

Reporting any information that is acquired as a direct result of research participation breaks confidentially by disclosing both the identity of the research participant(s) and the associated data. Thus, institutional ethics boards may require informed consent to include a statement that confidentiality will be kept *unless* collected data leads the researcher to reasonably suspect abuse or that the participant is a danger to themselves or others. Additionally, it is possible that a civil or criminal court might demand release of recordings and other data even when the researcher did not initiate the report and/or data is irrelevant to the court matter (cf. Khan, 2019). Disclosure of the researchers' responsibilities during informed consent may cause some participants to withdraw or provide inaccurate or misleading data to hide abuse, but these are necessary risks in order to properly convey the limits to confidentiality.

An additional consideration is the elapsed time between data collection and discovery of a reportable event. Frequently, data is not carefully scrutinized until well after collection, even years later. In such situations, the length of time that has passed since a particular incident was recorded may factor into whether there is a current risk to participants or other individuals, whether an illegal activity needs to be reported (e.g., if an activity that is currently illegal was legal at the time of the recording), and/or to what extent the parties involved are identifiable. Typically researchers do not have

the training to make such judgments for themselves and should consult their institutional ethics board and/or legal office for guidelines.

When considering whether and how to respond to potentially reportable activity, the ideal course of action is to set forth a detailed plan *before* data collection begins. This should be done in consultation with the relevant ethical review board, which may have certain policies already in place. Researchers should first ascertain whether they are mandated reporters, as defined by legal and employment contracts, and determine what qualifies as a reportable event in consideration of applicable laws, time since data collection, and the purported risk to participants or others. This process should also include understanding the procedure for reporting; for example, identifying the proper person or agency to contact, which may be a social agency or a designated party at the research institution, and also delineating how much information is necessary to disclose. Researchers should then specify how they plan to identify reportable events (e.g., only through incidental findings, or oversight on all of or a random sample of the recordings). Finally, researchers should establish how to incorporate information about mandated reporting, qualifying events, and any other risk to confidentiality into the informed consent, and whether/how to notify participants if reporting should occur. If no a priori plan was established and a potentially reportable event is identified, researchers should immediately contact their ethics review board to determine the best course of action.

In building this plan, it is important to remember that the "frontline" researchers are often student research assistants who may come from very different backgrounds and experiences to those of the population under study. This mismatch has the potential to lead to misunderstanding or misinterpretation of statements or actions on the part of the participants and different ideas about what is "normal". This mismatch should be explicitly addressed in training and policy (see section 4.3). In addition, while respecting legal requirements, it is also important to acknowledge that the impact of reporting is not equally distributed across populations; while all individuals are entitled to the protections of reporting (e.g., all children deserve equal protection from abuse), the actions of authorities may lead to harm to specific communities both historically (e.g., the "Sixties Scoop" on Canada's Indigenous populations) and at present.

Appropriate training of research staff

An important aspect in the ethical treatment of participant data is to provide appropriate training for research staff who will be working with participants, both in person (e.g., conducting informed consent and instructing on data collection procedures) and in post hoc data processing (e.g., listening to/ coding identifiable portions of audio recordings). Research assistants should be informed of how their human subjects training relates to specific issues regarding longform audio recordings, such as respectful and culturally informed discussion of language taken out of context, and keeping discussion about the recordings limited to other trained members of the research group.

If the project involves the study of any population not familiar to research assistants, study staff should receive cultural training on how typical and acceptable behaviors and social conventions differ across communities. This training step helps ensure that assistants do not misinterpret conversations or interactions on the recordings as dangerous or illicit. For example, physical punishment and drug use, while a potential concern in some societies, may be legal and socially acceptable in other societies. Principal investigators are responsible for establishing clear policies and criteria for determining whether an occurrence is a reportable event and for the marking of sensitive/protected material. However, if at any point a determination is unclear, researchers should seek the decision of a project supervisor for a final determination.

The level and type of training can differ depending on the extent to which research staff will have contact with identifiable information, or the extent to which the cultural background of the research assistants differs from the community under study. For example, for unidentifiable portions of audio recordings, such as extremely short snippets less than a second in length (Seidl, Warlaumont, & Cristia, 2019), extensive ethics training may not be required as there is little risk to participant confidentiality. Finally, research assistants should be informed that if listening to any recording is uncomfortable for them (e.g., threats of violence may trigger negative emotional responses), that they have the right to decline listening to audio, and appropriate protocols should be in place for whomever will continue working with that recording. In sum, principal investigators should consider risks both to participants/communities and to researchers themselves when designing appropriate staff training protocols.

Data sharing and storage

Because verification and replication are key to the scientific process, researchers should make a plan for data sharing prior to starting data collection. This data could include the quantified metrics, transcripts, or the audio recordings themselves in the form of small snippets or entire longform recordings. Most ethics boards are new to the complex issues surrounding longform recording research and data. Ethics boards will therefore vary dramatically in the strictness and care with which they approach issues of data storage and sharing, although many boards' practices around these issues are becoming more formalized as the issues surrounding big data become more salient.

Here we outline some important considerations when deciding whether and how to share longform audio recordings, followed by a brief list of potential repositories. We start from the assumptions that (a) data should be shared to the extent ethically possible, (b) data should be deleted only in rare circumstances, and (c) vulnerable populations require additional care in data sharing. Meyer (2018) provides a good general-purpose guide to data sharing and outlines a number of important pitfalls to avoid, such as promising to destroy or not to share data. The most important tips to keep in mind are as follows:

- **Avoid making promises you cannot (or should not) keep.** Often, researchers feel that their job is to mollify an ethics board. Researchers may be provided with templates for consent forms that are geared toward projects with less complex data-sharing concerns. These documents have historically contained promises to make data anonymous, to keep data sequestered in siloed laboratories, or to destroy data after a certain period of time, as well as limitations on how data will be used. It is important to view these templates as a suggestion and modify them when necessary or desirable (e.g., as data storage and sharing practices change over time). In our experience, ethics boards welcome this genre of thoughtful editing when it is carefully justified and explained. In the Ethics Repository associated with this article, we have included sample consent forms and scripts that we have employed in our own research (Cychosz et al., 2020).

- **Make an explicit plan for the long- and short-term storage and use of your data.** Our data live on well beyond the immediate time frame of the original study, and there is a growing recognition in the behavioral sciences of the benefits of data sharing (see section 3.1 above). For effective ethics and consent processes, it is important to think through when and how the data will be stored and shared. Selecting an appropriate data repository (e.g., Databrary [Databrary, 2012]; HomeBank [VanDam et al., 2016]) is an important component of this decision-making process, as well as who will have access to the data from the time they are collected to years or decades into the future. Some ethics applications have this perspective built in. If a given institution's ethics board does not take this approach, it is important to consider these decisions from the inception of the research project. How will the data get from the collection site to the lab and from the lab to the long-term data repository? Who will have access at each stage? What are the technical specifications necessary to ensure a sufficient level of protection against accidental data leakage? Who are the decision-makers regarding data sharing in the short and long term? What happens when the primary researcher dies or becomes unable to make decisions about changes to storage and sharing?

- **Consider the different kinds of data that will emerge from your recordings in deciding what can be shared, with whom, how, and for what purposes.** Different kinds of access may be granted to the raw longform recordings, vetted portions, transcripts, partial annotations, vocalizations shorter than 1 second, demographic information, metadata, and derived data, etc. (see Fig. 1). For example, rather than sharing the full audio recording, separate consents can be obtained to openly share transcripts of short segments that have been fully vetted for sensitive information and anonymized (MacWhinney, 2000). Similarly, many participants are open to sharing their recordings with the research community via password-protected systems, but would not wish to see them publicly shared on the Web.

- **Keep it simple for participants.** Before obtaining consent from participants, researchers should develop a basic plan for data storage and sharing based on the above principles. What is presented to the ethics board may be fairly complex, but things can be kept simpler for the participants. The fewer details you provide in the consent form, the less likely that some reasonable or necessary change in your storage and sharing system (which is almost certain to come up at some point down the road) will conflict with a promise made to the research participants. Do spell out clearly, in general terms, who will be allowed access to the data and under what conditions. If you provide options or levels of sharing, make sure they align with the options available in your planned data repository, such as public, semipublic, or private.

- **Don't forget meta-data.** A potential way to minimize the risks associated with data reuse that may not comply with participants' wishes (or participants' communities' wishes) is to limit the use of information about group-level identity for any research question outside those in the original study. Note, however, that researchers are sometimes obligated to report this information (e.g., reporting racial/ethnic groups in the USA), and it may not be possible to reasonably disseminate the findings without sharing this group-level information. Perhaps a better option is for the data to always be accompanied by a meta-data statement that explains the conditions for reuse, including contact information for the community whose data are represented, and any ethical guidelines this community provides.

Data repositories

Careful consideration of how collected data will be archived and accessed by other researchers and participant community members is crucial to ensuring that the participating communities have a chance of benefiting from the scientific work. There is no one-size-fits-all solution to ensuring that the

community benefits from archival data. While expectations for consent and privacy need to be met (see section 4.5 on first- and 4.6 on third-party informed consent), some level of sharing is imperative for scientific advancement and is *essential* when the documented linguistic and cultural data could serve as a cultural lifeline to communities whose traditions are under threat. Decisions about how to maximize value to the participating community should be done in advance of data collection so that informed consent can be appropriately conducted.

There are numerous options available for storage and sharing of data in various forms. Below is a partial list of repositories that are particularly well-suited for longform audio recordings.

Databrary (https://nyu.databrary.org) Databrary is a data repository for developmental research geared primarily toward video footage, but which can also be used for audio-only recordings. At the time of writing, it is not well-designed for easy interfacing with batch usage of a large number of files, but it is well-designed for smaller-scale projects and includes embedded structure for participant and file meta-data. Databrary requires an institution-to-institution legal contract. Once this is in place, authorized investigators can add research assistants as "affiliates" and manage different levels of access for affiliates and collaborators. There are multiple levels of access available for data sharing. At the time of writing, these include "private" (shared only within the lab – a good option for long-term storage of highly sensitive files or short-term storage of unvetted files that may be released to a larger audience at a later stage), "authorized users" (available to all authorized Databrary users, but may not be further distributed), "learning audiences" (available to authorized Databrary users and brief sections may be used for illustration purposes – note that this may include redistribution in public forums), or fully public.

HomeBank (https://homebank.talkbank.org/) HomeBank is a member of the TalkBank system. Contributors and users must have ethics training to apply for membership. After completing a brief phone interview, a contract is signed specifying terms of use. HomeBank was originally designed specifically for longform child language audio recordings and contains audio files, derivative data files, and transcripts, but could be extended to similar data sets centered on adult participants. At the time of writing, corpora may be shared according to four levels of access: fully public, members-only, secure (which requires additional approval, for more sensitive corpora), and embargoed (temporarily unavailable to all researchers outside of the primary team). Additional restrictions on use may be specified by the data contributor.

Institution-specific repositories Many institutions now provide their own resources for data storage and sharing that may be used for longform audio recordings, sometimes through existing systems such as Dataverse (Castro & Garnett, 2014). These will often have different levels of sharing suitable for sensitive and confidential files. Institution-specific repositories may also be ideal for promoting connections with speech communities interested in documenting/revitalizing their language(s) (e.g., AILLA, , 2017; CLA, , 2018). However, some institutional repositories may not have the user-vetting functions available through systems like Databrary or HomeBank.

Informed consent of participants

Longform audio recordings present unique challenges for obtaining informed consent. Following the principle of autonomy, participants need to understand what the recorders can capture as well as the myriad ways that the data can be used. Table 3 provides a summary of information that should be communicated to participants during the informed consent process. This includes details about the length and scope of the recording, the way in which the recording will be processed, and the types of data that will be extracted and analyzed (as overviewed in Fig. 1). A recent study suggests that parents' willingness to collect and share data from children's everyday lives varies in relation to the type of data collected and the level of sharing, suggesting these are important considerations for participants that should be explained clearly during the informed consent process (Levin et al., 2019).

Researchers may want to incorporate strategies to enhance (and check) comprehension during the informed consent process, such as (1) providing participants examples of the kind of data researchers will obtain from the recordings (e.g., an example audio snippet, distorted recording, quantitative output) or (2) using consent comprehension checks to ensure that participants understand the details of the data collection method (e.g., Lunshof, Chadwick, Vorhaus, & Church, 2008). As an example of a comprehension check, a researcher may explain orally how to pause the recorder, or turn it off, and may demonstrate how to do so with an example recorder. Then, after explaining the rest of the research method, the researcher may return to the topic of pausing/stopping and ask participants how to pause or stop the recorder to verify that they understood and remembered the original instructions. Further examples of consent comprehension checks like this are available in the informed consent scripts available in the OSF Ethics Repository associated with this paper.

Information about how recordings will be processed is likely to influence participants' calculation of the risks and benefits of the research and their decision to participate. For example, if researchers explain that recordings will only be processed by software to produce automated word counts, participants may have fewer privacy concerns and be more inclined to participate than if they believe human annotators will listen

to large swaths of their conversations. Many researchers may want to retain the option of listening to the audio recordings in the future, even if this is not part of their immediate research goals. Thus, unless there is a compelling reason not to, we recommend always asking participants for permission to listen to the recordings, and simultaneously including an alternative to opt out (e.g., "We may listen to the recordings you made *unless* you indicate that you do *not* want us to listen to the recording."). Alternatively, if the recordings are deemed particularly sensitive, listening to the recordings could be presented as an *opt-in* rather than an *opt-out* (e.g., "We will delete the recordings after they have been processed by our automated software *unless* you give us your explicit permission to retain the recordings and listen to them for additional analyses."). The options to opt-in or opt-out of allowing researchers to listen to recordings could be presented during the initial informed consent process (i.e., when participants enroll for the study) or in an additional consent form after participants have completed the recording. This latter strategy is preferable in many cases because it enables participants to have an idea of the content of the recording before deciding whether to allow researchers to listen to them. Example consent forms with opt-in/opt-out options are available in the OSF Ethics Repository associated with this paper.

Information conveyed during the informed consent process could also affect participants' behavior when conducting the recording. For example, a parent who believes researchers are interested in their child's language development may engage in activities that showcase their child's language, and may pay less attention to other family dynamics that are not the focus of the research. Although researchers may be tempted to hide the goal of the study so as to not alter participants' behaviors, disclosing the study's goals is an important element of respecting participants' autonomy.

However, a challenging issue relates to how to inform participants about all the possible uses of the data, including ones that the researcher may not have yet considered. Given the large number of secondary research questions that could be asked from longform recordings, it may be hard to include all future uses of the data in the consent form. One solution is to offer tiered consent options that give participants control over how broadly their data can be used for secondary research purposes (Meyer, 2018). Participants could choose the broadest option, allowing reuse of the data for any potential research question, or they could choose to set limits on the kinds of secondary analyses that may be conducted. For example, if the original study's aim is to investigate the relationship between the home environment and child language development, one option might let participants consent to the data being used to study other aspects of child development while limiting other uses of the recordings (e.g., marital conflict). However, if such restrictions are provided as an option, the researcher must have a plan to ensure compliance. Even

for data repositories that allow individualized restrictions on use (not all do), there is the risk that secondary data users may (accidentally) fail to abide by the restrictions.

Third-party consent

In addition to participants who actively consent to the recording, additional third-party members often incidentally appear on real-world longform audio recordings. These may be family members living in the same household, friends who are visiting, and even other people who come within the range of the recording device. These people in principle do not have direct communication with the researcher, and yet should be covered by the same guiding principles of autonomy, beneficence, justice, and privacy/confidentiality.

What to do with this inevitable consequence of naturalistic audio recordings? First, researchers should be familiar with the laws governing audio recordings in the region in which they are practicing. For example, a majority of states in the U.S. only require one-party consent to audio-record (Anderson, 2017). However, 12 U.S. states, including California and Illinois, require all-party consent (often referred to as two-party consent) for audio recording. This means that each individual appearing in a recording must consent to it. Interpretation of all-party consent laws differ by state and country. Canada is unilaterally a one-party consent jurisdiction (Criminal Code RSC, 1985), while Germany is all-party consent (Strafgesetzbuch, n.d.), and South Korea is one-party so long as the recorder is a member of the conversation (Tongsinbimilbohobeob, 1993). As a result of these differences between countries and states/provinces, we encourage researchers to consult a legal resource, such as a university's in-house counsel, to ensure that they are complying with local law (Manson & Robbins, 2017; Robbins, 2017). This is especially relevant if participants cross boundaries/borders between states or countries. For the United States, to determine where one's university/state falls legally, researchers may also consult legal summary publications such as Matthiesen and Lehrer (2019), which straightforwardly summarizes recording consent processes in each of the United States.

Exceptions to these stringent policies are made when participants cannot be said to have a "reasonable expectation of privacy", such as in a public arena (Fogel, 2014; Robbins, 2017). A gray area arises for third parties in the home. Current legal debates surrounding the recording ethics of Amazon's Alexa technology provide some clarification (Manning, 2019). Alexa is an assistive device for the home, activated with verbal key words (e.g., "Hey, Alexa … "). Following this activation, the device records the subsequent interaction which is then stored on Amazon's servers. The Alexa device owner obviously does not then have a reasonable expectation of privacy, as they have brought this technology into their home, but the unknowing third party likely still

does. Crucially for our purposes, the onus for two-party consent, that is to ensure that all recorded parties are made aware, does not fall on Amazon. Instead, the Alexa owner should inform all visitors of the recording device.

This suggests several nontechnical ways to mitigate third-party concerns. In the case of data collection in the home, or other constrained locations like a preschool or daycare, the easiest path may be to receive verbal permission from all household members or individuals such as neighbors who are likely to appear on the recording, even those who may appear sparingly. Likewise, researchers should encourage participants to inform their visiting friends and neighbors that they or their child are wearing a recording device. To reduce participant burden in literate contexts, researchers can provide an information card that explains the study and the recording device, which can be given to any individual who may appear in the recording. This information card can let third parties know of any available options to delete sensitive or compromising interactions from the recordings. Although uncommon, some ethics boards may require researchers to verify that third parties received this information, for example by asking parties to sign the information card.

In situations where unanticipated individuals might appear, such as a public market or a schoolyard, it may be possible to post a sign indicating active recording with a place to give oral/sign consent before entry. Another option to mitigate third-party concerns is for participants to wear colorful, distinctive clothing so that uninterested parties can work or socialize elsewhere. However, this may also come at a cost, for both participants and researchers. Distinctive clothing can result in a more self-monitored style of interaction. When the target participant is a child, using distinctive clothing may be problematic because it makes them visibly different from (and perhaps teased by) their peers during the recording process.

As a final option, researchers can attempt to remove third-party data from the longform recordings (e.g., prior to sharing the recordings). However, this process is time-consuming and often impossible in the case of overlapping or noisy speech environments. Additionally, there may be a significant cost on the validity of the data for many research questions.

Finally, researchers must bear in mind the civil liberty structure of the country in which they are doing research. While in the U.S., for example, Fourth Amendment rights guarantee protection from undue search and seizure (right to privacy), researchers should not assume that these liberties are ubiquitously guaranteed, particularly in regions with more authoritarian governments. In our highly connected world, concepts of privacy are changing, and legal precedent is emergent. As much as social science research should, ideally, be international, our research methods must be on solid legal and cultural ground for the country or countries in question. To the extent that it is possible, special care should be taken to ensure that all members of marginalized groups understand the power

and perpetuity of the recordings, even if it means that the research team must make compromises (e.g., only recording individuals in the home and not at public meeting spaces or large, multiparty contexts).

Next steps and conclusion

In this article, we outlined ethical issues surrounding research using longform recordings of everyday life, paying particular attention to using this methodology with vulnerable populations who are underrepresented in social science research. We first summarized ethical principles that guide social scientists, including participant autonomy, respect for participant privacy/confidentiality, and justice. We then outlined the benefits and risks that longform recording methodologies can bring for participants, communities, and researchers. Finally, we proposed best practices to address ethical challenges, ranging from how to ensure the confidentiality of *all* individuals who appear on recordings to how to handle illicit activities on recordings. To that end, we also direct readers to the Ethics Repository associated with this article (Cychosz et al., 2020) where they can find sample consent forms and scripts, links to other relevant resources such as the vetting guide we discussed, and participant frequently asked questions.

For readers interested in incorporating longform recordings into their own research, we encourage a few practical steps. First, readers may wish to participate in the Daylong Audio Recordings of Children's Linguistic Environments (DARCLE) community (darcle.org) where new, unpublished research is often shared and ethical issues surrounding this methodology are often discussed. Even for those not interested in child development or language use, the DARCLE group can provide essential support and feedback for a variety of research pursuits in the behavioral and social sciences. Second, we direct readers to a couple of additional publications that complement topics addressed here, notably the comprehensive guide to longform audio collection in Casillas and Cristia (2019) and the HomeBank recording vetting manual (VanDam et al., 2018), available for download through the DARCLE website.

Technological advances in recording methodologies afford contemporary researchers many new and exciting avenues for investigation. However, data methodologies such as longform recordings of everyday life may also bring increased risk to participants, communities, and members of the research team. Researchers employing this methodology can implement a range of best practices for responsibly collecting, storing, and sharing these data to ensure maximal beneficence to both participants and the broader scientific community.

Open practice statement This article did not make use of any data sets, and no experiments were performed.

Author contributions MCy, RR, and AW planned and drafted the paper, designed the figure and tables, and reconciled and integrated feedback; all authors discussed the topics, wrote sections, and edited the paper.

References

(AILLA). Archive of the Indigenous Languages of Latin America. (2017). *Archive of the Indigenous Languages of Latin America*. University of Texas, Austin. https://ailla.utexas.org.

Alharbi, R., Stump, T., Vafaie, N., Pfammatter, A., Spring, B., & Alshurafa, N. (2018). I can't be myself: Effects of wearable cameras on the capture of authentic behavior in the wild. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *2*(3), 1–40. https://doi.org/10.1145/3264900

Anderson, A. (2017, Winter). Classroom Taping Under Legal Scrutiny-A Road Map for a Law School Policy. *Journal of Legal Education, 66*(2). 372-408.

Anderson & Fausey (2019). Modeling nonuniformities in infants' everyday speech environments. Presented at the Biennial Meeting of the Society for Research on Child Development. Baltimore, MD.

Arnett, J. J. (2008). The neglected 95%: Why American psychology needs to become less American. *The American Psychologist*, *63*(7), 602–614.

Avineri, N., Blum, S., Johnson, E. J., Zentella, A. C., Brice-Heath, S., McCarty, T., Rosa, J., Flores, N., Ochs, E., Samy Alim, H., Paris, D. & Kremer-Sadlik, T. (2015). Invited Forum: Bridging the "Language Gap". *Linguistic Anthropology*, *25*(1), 66–86. https://doi.org/10.1111/jola.12071.6

Brenzinger, M., & Heinrich, P. (2013). The return of Hawaiian: Language networks of the revival movement. *Current Issues in Language Planning*, *14*(2), 300–316.

Bucher, R., Fritz, C., & Quarantelli, E. (1956). Tape recorded interviews in social research. *American Sociological Review, 21*(3), 359-364.

California Language Archive (CLA). (2018). California Language Archive. University of California, Berkeley. http://cla.berkeley.edu.

Canault, M., Le Normand, M.-T., Foudil, S., Loundon, N., & Thai-Van, H. (2016). Reliability of the Language ENvironment Analysis system (LENA^TM) in European French. *Behavior Research Methods*, *48*(3), 1109–1124.

Casillas, M., Brown, P., & Levinson, S. C. (2019). Early language experience in a Tseltal Mayan village. *Child Development*, cdev.13349.

Casillas, M., and Cristia, A. (2019). A step-by-step guide to collecting and analyzing long-format speech environment (LFSE) recordings. *Collabra: Psychology*, *5*(1), 24.

Castro, E., & Garnett, A. (2014). Building a bridge between journal articles and research data: The PKP-Dataverse integration project. *International Journal of Digital Curation*, *9*, 176-184.

Chen, F., Adcock, J., & Krishnagiri, S. (2008) Audio privacy: reducing speech intelligibility while preserving environmental sounds. In *Proceedings of the 16th ACM international conference on Multimedia*, 733–736.

Chen, Z., Klatzky, R., Siewiorek, D., Satyanarayanan, M., Hu, W., Wang, J., … & Pillai, P. (2017). An empirical study of latency in an emerging class of edge computing applications for wearable cognitive assistance. *Proceedings of the Second ACM/IEEE Symposium on Edge Computing*, *17*, 1-14.

Criminal Code RSC. (1985). c. C - 46, s. 184.

Cristia, A., Dupoux, E., Gurven, M., & Stieglitz, J. (2017). Child-Directed speech is infrequent in a forager-farmer population: A time allocation study. *Child Development*, *90*(3), 759–773.

Cristia, A., Ganesh, S., Casillas, M., & Ganapathy, S. (2018). Talker diarization in the wild: The case of child-centered daylong audio-recordings. In *Proceedings of Interspeech 2018* (pp. 2583-2587). doi:https://doi.org/10.21437/Interspeech.2018-2078.

Cychosz, M., Cristia, A, Bergelson, E., Casillas, M., Baudet, G., Warlaumont, A. S., Scaff, C., Yankowitz, L., & Seidl, A. (2019). *BabbleCor: A Crosslinguistic Corpus of Babble Development in Five Languages*. https://doi.org/10.17605/OSF.IO/RZ4TX

Cychosz, M., Romeo, R., Soderstrom, M., Scaff, C., Ganek, H., Cristia, A., Casillas, M., de Barbaro, K., Bang, J., & Weisleder, A. (2020). *Daylong Audio Recording of Children's Linguistic Environments (DARCLE) Ethics Repository*. Open Science Framework Project: https://osf.io/u3tfv/.

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

de Barbaro, K. (2019). Automated sensing of daily activity: A new lens into development. *Developmental Psychobiology*, *61*(3), 444-464.

Department of Health and Human Services. (2015). *Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule*. Retrieved October 1, 2019 from https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html .

Derry, S. J., Pea, R. D., Barron, B., Engle, R. A., Erickson, F., Goldman, R., … & Sherin, B. L. (2010). Conducting Video Research in the Learning Sciences: Guidance on Selection, Analysis, Technology, and Ethics. *Journal of the Learning Sciences*, *19*(1), 3–53.

Dubowitz, H., Hein, H., & Tummala, P. (2018). World perspectives on child abuse, 13th Ed. Aurora, CO: International Society for the Prevention of Child Abuse and Neglect.

Fasick, F. (1977). Some uses of untranscribed tape recordings in survey research. *The Public Opinion Quarterly, 41*(4), 549-552.

Ferjan Ramírez, N, Lytle, SR, Fish, M, Kuhl, PK. (2019). Parent coaching at 6 and 10 months improves language outcomes at 14 months: A randomized controlled trial. Developmental Science, *22*(3), e12762.

Fogel, J. (2014) From the bench: A Reasonable expectation of privacy. *The American Bar Association Litigation Journal.* Retrieved March 15, 2019 from https://www.americanbar.org/groups/litigation/publications/litigation_journal/2013-14/spring/a_reasonable_expectation_privacy/.

Ganek, H., Smyth, R., Nixon, S., & Eriks-Brophy, A. (2018). Using the Language ENvironment Analysis (LENA) System to Investigate Cultural Differences in Conversational Turn Count. *Journal of Speech, Language, and Hearing Research*, *61*(9), 2246–2258.

Garcia, M. (2016). Racist in the machine: The disturbing implications of algorithmic bias. *World Policy Journal, 33*(4), 111-117.

Gardner, F. (2014). Methodological issues in the direct observation of parent–child interaction: Do observational findings reflect the

natural behavior of participants? *Clinical Child and Family Psychology Review, 3*, 185–198. 10.1023/A

Gehlert, S., & Mozersky, J. (2018). Seeing beyond the margins: Challenges to informed inclusion of vulnerable populations in research. *The Journal of Law, Medicine & Ethics, 46*(1), 30–43.

Greenwood, C. R., Carta, J. J., Walker, D., Watson-Thompson, J., Gilkerson, J., Larson, A. L., & Schnitz, A. (2017). Conceptualizing a public health prevention intervention for bridging the 30 million word gap. *Clinical Child and Family Psychology Review, 20*(1), 1–22.

Harari, G. M., Müller, S. R., Aung, M. S., & Rentfrow, P. J. (2017). Smartphone sensing methods for studying behavior in everyday life. *Current Opinion in Behavioral Sciences, 18*, 83–90.

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

Hill, R. P. (1995). Researching sensitive topics in marketing: The special case of vulnerable populations. *Journal of Public Policy & Marketing, 14*(1), 143–148.

Kay, M., Choe, E. K., Shepherd, J., Greenstein, B., Watson, N., Consolvo, S., & Kientz, J. A. (2012). Lullaby. *Proceedings of the 2012 ACM Conference on Ubiquitous Computing – UbiComp,* 12.

Kelly, P., Marshall, S. J., Badland, H., Kerr, J., Oliver, M., Doherty, A. R., & Foster, C. (2013). An ethical framework for automated, wearable cameras in health behavior research. *American Journal of Preventive Medicine, 44*(3), 314–319.

Khan, S. (2019). The Subpoena of ethnographic data. *Sociological Forum, 34*(1), 253–263.

Kiukkonen, N.J.B., Dousse, O., & Gatica-Perez, D. J. (2010). Towards rich mobile phone data sets: Lausanne data collection campaign. In *Proceedings ACM International Conference on Pervasive Services* (ICPS2010).

Kumar, S., Nguyen, L. T., Zeng, M., Liu, K., & Zhang, J. (2015). Sound shredding: Privacy preserved audio sensing. In *Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications*, pp. 135-140.

Lane, N. D., Mohammod, M., Lin, M., Yang, X., Lu, H., Ali, S., Doryab , A., Berke, E., Choudhury, T., and Campbell, A. (2011) Bewell: A smartphone application to monitor, model and promote wellbeing. *PervasiveHealth*.

Langheinrich, M. (2001). Privacy by Design — Principles of Privacy-Aware Ubiquitous Systems. In G. D. Abowd, B. Brumitt, & S. Shafer (Eds.), Ubicomp 2001: Ubiquitous Computing (pp. 273–291). Springer Berlin Heidelberg.

Larson, E. C., Lee, T., Liu, S., Rosenfeld, M., and Patel, S. N. (2011). Accurate and privacy preserving cough sensing using a low-cost microphone. In *Proceedings of the13th international conference on Ubiquitous computing*, 375–384.

Le Franc, A., Riebling, E., Karadayi, J., Wang, Y., Scaff, C., Metze, F., & Cristia, A. (2018). The ACLEW DiViMe: An Easy-to-use Diarization Tool. In *Proceedings of Interspeech 2018*. (pp. 1383-1387).

Lee, R. (2004). Recording Technologies and the Interview in Sociology, 1920–2000. *Sociology, 38*(5). 869-889.

Leung, C. Y. Y., Hernandez, M. W., & Suskind, D. L. (2019). Enriching home language environment among families from low-SES backgrounds: A randomized controlled trial of a home visiting curriculum. *Early Childhood Research Quarterly*.

Levin, H., Egger, D, Johnson, M., & de Barbaro, K. (2019). Parent Willingness to Collect and Share Children's Mobile-Sensing Data. (under review). Available at: https://psyarxiv.com/u39xg.

Lunshof, J. E., Chadwick, R., Vorhaus, D. B., & Church, G. M. (2008). From genetic privacy to open consent. *Nature Reviews Genetics, 9*(5), 406–411.

Luo, R., Alper, R.M., PhD, CCC-SLP, Hirsh-Pasek, K., Mogul, M., Chen, Y., Masek, L. R., Paterson, S., Pace, A., Adamson, L. B., Bakeman, R., Golinkoff, R. M., Owen, M. T.. (2019).

Community-Based, Caregiver-Implemented Early Language Intervention in High-Risk Families: Lessons Learned. *Progress in Community Health Partnerships: Research, Education, and Action, 13*(3), 282-291.

MacWhinney, B. (2000) The CHILDES project: Tools for analyzing talk (3rd). Mahwah, NJ: Lawrence Erlbaum Associates.

Manning, G. (2019). Alexa: Can You Keep a Secret? The Third-Party Doctrine in the Age of the Smart Home. *American Criminal Law Review, 56*. Online.

Manson, J., & Robbins, M. L. (2017). New evaluation of the electronically activated recorder (EAR): Obtrusiveness, compliance, and participant self-selection effects. *Frontiers in Psychology, 8*(658), 1-9.

Matthiesen, W., & Lehrer, S.C. (2019). Laws on recording conversations in all 50 states. Available at https://www.mwl-law.com/wp-content/uploads/2018/02/RECORDING-CONVERSATIONS-CHART.pdf

Mehl, M. R. (2017). The Electronically Activated Recorder (EAR) A Method for the Naturalistic Observation of Daily Social Behavior. *Current Directions in Psychological Science, 26*(2), 184-190.

Mehl, M. R., & Pennebaker, J. W. (2003). The sounds of social life: A psychometric analysis of students' daily social environments and natural conversations. *Journal of Personality and Social Psychology, 84*(4), 857.7

Meyer, M. N. (2018). Practical tips for ethical data sharing. *Advances in Methods and Practices in Psychological Science, 1*(1), 131–144.

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., … Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour, 1*(1), 0021.

Nickel, P. J. (2006). Vulnerable populations in research: the case of the seriously ill. *Theoretical medicine and bioethics, 27*(3), 245-264.

Nielsen, M., Haun, D., Kartner, J., Legare, C.H. (2017) The persistent sampling bias in developmental psychology: A call to action. *Journal of Experimental Child Psychology, 162*, 31-38.

Office for Human Research Protections of the U.S. Department of Health & Human Services. (2019). Federal Policy for the Protection of Human Subjects ('Common Rule'). Retrieved from https://www.hhs.gov/ohrp/regulations-and-policy/regulations/common-rule/index.html

Räsänen, O., Seshadri, S., Karadayi, J., Riebling, E., Bunce, J., Cristia, A., Metze, F., Casillas, M. Rosemberg, C., Bergelson, E., Soderstrom, M. (2019). Automatic word count estimation from day-long child-centered recordings in various language environments using language-independent syllabification of speech. *Speech Communication, 113*, 63–80.

Robbins, M. L. (2017). Practical suggestions for legal and ethical concerns with social environment sampling methods. *Social Psychological and Personality Science 8*(5), 573–580.

Rogers, W., & Lange, M. M. (2013). Rethinking the vulnerability of minority populations in research. *American Journal of Public Health, 103*(12), 2141–2146.

Seidl, A., Warlaumont, A., & Cristia, A. (2019). Towards detection of canonical babbling by citizen scientists: Performance as a function of clip length. In *Proceedings of Interspeech 2019*. Graz, Austria.

Sieber, J. (1994). Issues presented by mandatory reporting requirements to researchers of child abuse and neglect. *Ethics & Behavior, 4*(1), 1-22.

Soderstrom, M., & Wittebolle, K. (2013) When do caregivers talk? The influences of activity and time of day on caregiver speech and child vocalizations in two childcare environments. PLOS ONE 8(11): e80646.

Stanley, T. D., Carter, E. C., & Doucouliagos, H. (2018). What meta-analyses reveal about the replicability of psychological research. *Psychological Bulletin, 144*(12), 1325-1346.

Strafgesetzbuch [StGB]. (n.d.). [Penal Code], § 201.

Talbert, E. (2019). Beyond data collection: Ethical issues in minority research. *Ethics and Behavior, 29*(7), 531–546.

Tamis-LeMonda, C. S., Kuchirko, Y. , Luo, R. , Escobar, K. and Bornstein, M. H. (2017), Power in methods: language to infants in

structured and naturalistic contexts. *Developmental Science*, 20: e12456.

Tongsinbimilbohobeob. (1993). [Protection of Communications Secrets Act], Act No. 4650.

VanDam, M., Warlaumont, A. S., Bergelson, E., Cristia, A., Soderstrom, M., De Palma, P., & MacWhinney, B. (2016). HomeBank: An online repository of daylong child-centered audio recordings. *Seminars in Speech and Language*, *37*(2), 128–142.

VanDam, M., Warlaumont, A.S., MacWhinney, B., Soderstrom, M., & Bergelson, E. (2018). Vetting Manual: Preparation of recordings for unrestricted publication in HomeBank Version 1.1. https://homebank.talkbank.org/pubs/vetting.pdf

Wang, R., Chen, F., Chen, Z., Li, T., Harari, G., Tignor, S., … & Campbell, A. T. (2014a). StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing*, 3-14.

Wang, Z., Pan, X., Miller, K. F., & Cortina, K. S. (2014b). Automatic classification of activities in classroom discourse. *Computers & Education*, 78, 115–123.

Weisleder, A., & Fernald, A. (2013). Talking to children Matters: early language experience strengthens processing and builds Vocabulary. *Psychological Science*, *24*(11), 2143–2152.

World Medical Association. (2013). Declaration of Helsinki: Ethical principles for medical research involving human subjects.

Wyatt, B., Choudhury, T., and Bilmes, J. (2007). Conversation Detection and Speaker Segmentation in Privacy-Sensitive Situated Speech Data. In *Proceedings of Interspeech 2007*.

Xu, D., Yapanel, U., & Gray, S. (2009). LENA TR-05: Reliability of the LENA Language ENvironment Analysis System in young children's natural home environment. Boulder, CO: LENA Foundation.

Zhang, Y., Xu, X., Jiang, F., Gilkerson, J., Xu, D., Richards, J., Harnsberger, J., Topping, K. (2015). Effects of quantitative linguistic feedback to caregivers of young children: A pilot study in China. *Communication Disorders Quarterly, 37*(1), 16-24.