Received: 7 November 2021 | Revised: 9 June 2022 | Accepted: 9 September 2022

DOI: 10.1002/icd.2377

METHODOLOGICAL ARTICLE

WILEY

Open science considerations for descriptive research in developmental science

Jessica E. Kosie | Casey Lew-Williams

Department of Psychology, Princeton University, Princeton, New Jersey, USA

Correspondence
Jessica E. Kosie, Department of Psychology, Princeton University, Princeton, New Jersey, USA.
Email: jkosie@princeton.edu

Funding information
Eunice Kennedy Shriver National Institute of Child Health and Human Development, Grant/Award Numbers: F32HD103439, R01HD095912

[Corrections added on 8 November 2022, after first online publication: "special population" has been corrected to "understudied population" in sections 5 and 7]

Abstract

Descriptive developmental research seeks to document, describe and analyze the conditions under which infants and children live and learn. Here, we articulate how open-science practices can be incorporated into descriptive research to increase its transparency, reliability and replicability. To date, most open-science practices have been oriented toward experimental rather than descriptive studies, and it can be unclear how to translate open-science practices (e.g., preregistration) for research that is more descriptive in nature. We discuss a number of unique considerations for descriptive developmental research, taking inspiration from existing open-science practices and providing examples from recent and ongoing studies. By embracing a scientific culture where descriptive research and open science coexist productively, developmental science will be better positioned to generate comprehensive theories of development and understand variability in development across communities and cultures.

KEYWORDS
child development, descriptive research, infant development, open science, transparency

Highlights

- Descriptive research is vital to the development of comprehensive theories of early development, and open-science practices can increase the transparency and replicability of this effort.
- Many existing open-science practices are oriented toward experimental rather than descriptive research.

- Considerations related to the adoption of open-science practices in descriptive research are discussed, complemented by examples of successful adoption.

1 | INTRODUCTION

Descriptive studies of infants' and children's behaviour and input are vital for theories of early learning and development. Though developmental science has been dominated by experimental designs for many years, descriptive research—which documents, describes, and analyses the conditions under which infants live and learn in a variety of natural settings—has had a continuous and vibrant presence, and has played a primary role in shaping our understanding of early development. Descriptive research provides important details about the natural contexts in which infants learn and, beyond its strong independent value, guides the design of more ecologically-valid experimental studies. In these ways, descriptive research generates new insights that would be missed by focusing on experiments alone.

Across all fields of psychology and related disciplines, there has been a shift in the past decade toward the wide adoption of practices that increase the openness, transparency, and replicability of research (Nelson et al., 2018; Spellman et al., 2018). Many of the developmental scientists who are incorporating descriptive studies into their own research programmes have also readily adopted open-science practices (e.g., sharing audio and video recordings alongside detailed coding guidelines and, in some cases, even coding software or scripts), and their work highlights the many benefits that open science offers for descriptive research. At the same time, however, incorporating open science practices into descriptive research can raise unique challenges. Thus, adoption of practices that increase the transparency of descriptive work also requires researchers to carefully balance the costs and benefits of incorporating open-science practices.

Throughout this paper, we discuss how descriptive research and open science can coexist productively, inspired by existing open-science practices but geared toward helping those who intend to do open, descriptive research in the field of infant and child development. In an effort to understand and describe the varied dynamics of infants' everyday learning environments, we have increased the use of descriptive methods in our own research over recent years (e.g., Casey et al., in prep; Kosie & Lew-Williams, in prep). Transparency and replicability are of primary concern, we have drawn inspiration from the practices of other researchers who are doing related descriptive work in a way that is open and transparent (e.g., Bergelson, Amatuni, et al., 2019; Bergelson, Casillas, et al., 2019; Bunce et al., 2020; Casillas et al., 2020; Cychosz et al., 2021; Herzberg et al., 2021; Hoch et al., 2019, 2020; Karasik et al., 2018; Kremin et al., 2020; Mendoza & Fausey, 2021a, 2021b; Romeo et al., 2018; Soderstrom et al., 2021; Sullivan et al., 2020; Warlaumont et al., 2014).

Our goal is to share this accumulated knowledge with others, highlighting examples and making suggestions for how open-science practices can be incorporated into descriptive research. First, we clarify what we mean by "descriptive research" for the purposes of this article. We then discuss unique challenges to descriptive, as compared to experimental, research. Next, we outline a set of considerations related to these challenges, focusing on suggestions for increasing transparency and reproducibility. Throughout the paper, we pose a number of open questions and concerns that we hope will spark discussion about how to best continue promoting and increasing transparency in descriptive developmental science.

2 | WHAT DO WE MEAN BY "DESCRIPTIVE RESEARCH"?

The category of "descriptive research" is wide-ranging. It can include qualitative, quantitative, and mixed-methods studies across a variety of behavioural disciplines, including psychology, linguistics, education, sociology and anthropology (to name just a few). Descriptive research in the behavioural sciences often involves observing events—either via live observation or audio/video recordings—and annotating different aspects of situations as they naturally unfold. These descriptions of events are then organized in a way that enables a reader to understand various features of the observed activity (e.g., the occurrence and timing of behaviours), and frequently include data visualization or summary statistics across individuals or groups to clearly communicate the effects that may be present.

Early work in the field of developmental psychology was largely descriptive in nature, often involving individuals' detailed observations of infants' and children's physical, motor and language development (e.g., Darwin, 1877; Gesell, 1928, 1940; McGraw, 1935; Preyer, 1888; Shinn, 1900). In the mid-1900s, the rise of behaviourism and the development of new methods to index infants' cognitive processing—for example, methods based on infant looking (see Aslin, 2007 for a review)—led to an increase in experimental studies of early development. Experimental research quickly became the dominant mode of investigation. Despite this prioritization of experimental work, descriptive research has continued to be a central method for investigating early development and has fostered a greater understanding of the natural conditions in which infants and children live and learn. The independent value of descriptive work has been evidenced by both early and more contemporary research. For example, in the 1920s and 1930s, Arnold Gesell developed new methods for observing and documenting early development, which resulted in a catalogue of infant and child behaviour and the construction of a set of developmental norms that—for better or worse (see Thelen & Adolph, 1992 for further discussion)—are still used today (e.g., Gesell, 1928; Gesell & Ilg, 1943; Gesell & Thompson, 1934). A second example comes from more contemporary work on infants' natural walking behaviours. This research provides insights into the development of infants' walking ability and how the shift from crawling to walking shapes infants' input from the world and from other people (e.g., Adolph et al., 2012; Adolph & Tamis-LeMonda, 2014; Karasik et al., 2011; Karasik et al., 2014). This more comprehensive understanding of motor skill development—grounded in descriptive methods—has generated clinical implications that extend beyond motor skill development (e.g., Adolph et al., 2018). Additionally, recent descriptive work has revealed marked differences between the way that learning is typically tested in the lab and how it occurs in infants' everyday lives (Casey et al., in prep; Kosie & Lew-Williams, in prep; Clerkin et al., 2017; Fausey et al., 2016). Descriptive studies of infants' everyday environments provide insight into the mechanisms that underlie early learning, help generate informed hypotheses, enable testing of hypotheses in ways that are ecologically valid, allow for comparison of in-lab and at-home infant behaviours and contribute to comprehensive theories of infant development. As the field moves toward conducting more cross-cultural and cross-community research, there may be practical and intellectual advantages to prioritizing descriptive research.

For the purposes of the current paper, we restrict our definition of "descriptive research" to the types of studies we most commonly design and encounter in our area of developmental psychology—early cognitive development—and to research that is more quantitative in nature (e.g., involving numerical summaries and statistical analyses). However, qualitative researchers have thought carefully about open science considerations (e.g., Braun & Clarke, 2021; Haven et al., 2020; Humphreys et al., 2021), and there is much to learn from their work, though we do not directly address qualitative research in the current paper. Further, while much neuroscientific research in this area is often descriptive as well (e.g., physiological mapping studies), our focus is on behavioural studies.

Descriptive studies in early cognitive development frequently seek to characterize the observed behaviours of infants, children and caregivers, either individually or during naturalistic interactions. Examples of this type of descriptive research include investigations of the dynamics of caregiver-infant interactions (e.g., Kosie & Lew-Williams, in prep; Custode & Tamis-LeMonda, 2020; Romeo et al., 2018; Rowe et al., 2008; Tincoff et al., 2019), speech available in infants' and children's everyday environment (e.g., Casey et al., in prep; Bergelson, Amatuni, et al., 2019; Bergelson, Casillas, et al., 2019; Casillas et al., 2020; Kremin et al., 2020; Roy et al., 2015; Weisleder & Fernald, 2013), infants' everyday musical experience (Mendoza & Fausey, 2021a), the prevalence of varied categories of visual information in infants' everyday visual input (Clerkin et al., 2017; Fausey et al., 2016), and more. While some studies are purely descriptive, with the focused goal of describing behaviour, many other studies involve first describing behaviour and then linking these observations to each other or to other measures (e.g., Romeo et al., 2018; Rowe et al., 2008; Weisleder & Fernald, 2013).

3 | DESCRIPTIVE WORK POSES UNIQUE CHALLENGES FOR OPENNESS AND TRANSPARENCY

Researchers studying infant and child development have been responsive to the field's unique challenges in addressing issues related to open science (see Davis-Kean & Ellis, 2019; Eason et al., 2017; Gennetian et al., 2022; Peterson, 2016 for further discussion of these issues) and have led the field toward increasing transparency and replicability in a variety of ways. These responses include developmental-specific considerations about preregistration (Havron et al., 2020), improving measurement (Byers-Heinlein et al., 2021; DeBolt et al., 2020), thinking carefully about sample size issues (Oakes, 2017; Schott et al., 2019), sharing video, audio, and other data (Cychosz et al., 2020; Gilmore et al., 2020, 2021), meta-analysis (Bergmann et al., 2018), and collaborative research (Byers-Heinlein et al., 2020; Frank et al., 2017). While these varied responses to replicability issues apply more or less well to a broad range of developmental research, they have primarily been oriented toward experimental rather than descriptive studies. Existing open-science principles translate less directly to work that is more descriptive in nature, though this domain of research is an essential part of developmental science—and in the future, we hope it takes on a central role.

In infancy research, experimental studies frequently involve exposing infants to carefully constructed tasks and measuring their behaviour. For example, researchers are commonly interested in the location and duration of infant looks, patterns of neural activity, or overt motor behaviours (e.g., pointing, reaching for objects, exchanging objects with another person). Experiments are almost always designed with the goal of eliciting a particular behaviour and drawing inferences based on the sample of that behaviour (e.g., using infant looking as an index of attention or even "preference"). Researchers almost always have a hypothesis about what patterns of behaviour to expect under different conditions and can frequently generate some idea of the size of the effect. These features of experimental work lend themselves well to open-science practices like preregistration, and the numerical data that results from many experimental studies can easily be de-identified for sharing. While this does not necessarily make it "easy" to do open, replicable science, it does make it more straightforward to implement existing open-science practices into each phase of a project.

Descriptive research, on the other hand, differs from experiments across key dimensions. Existing "best practices" for open science do not translate clearly too descriptive studies, and descriptive research may be more vulnerable to issues related to transparency and reproducibility. For example, instead of having strong expectations about a restricted set of behaviours, researchers doing descriptive work may first want to observe all that is occurring during natural interactions, and only then classify the behaviours of infants and caregivers (e.g., different categories of gestures or words). In such cases, researchers may not start out with clear, pre-existing expectations about the behaviours that will occur across individuals and contexts (e.g., Prather, 2021).

Contemporary researchers doing descriptive work have served as an example of how to address these issues. Inspired by their practices, we now turn to describe key considerations for conducting open-ended descriptive research. In turn, we discuss preregistration, sample size justification, open procedures for coding, and sharing audio/video data.

4 | PREREGISTERING DESCRIPTIVE RESEARCH

Preregistration involves specifying—often in advance of data collection—hypotheses and predictions, information about procedures and materials, variables, sampling plans, exclusion criteria, and analysis plans with the goal of decreasing the occurrence of false positives and "researcher degrees of freedom" across all phases of experimental design and implementation (Simmons et al., 2011, 2021). In many ways, the objectives of preregistration seem to contradict the goals of descriptive research, which is often considered to be primarily exploratory (e.g., looking for stable patterns in the absence of a priori predictions; Scheel et al., 2021). In practice, though, much descriptive work in our field often involves pre-planned analyses, such as linking features of infant-directed input to behavioural and neurological measures of language development (e.g., Kosie & Lew-Williams, in prep; Romeo et al., 2018; Weisleder & Fernald, 2013) or comparing the frequency of different classifications of behaviour (e.g., between- versus within-sentence language switching in bilingual infants' everyday language input; Kremin et al., 2020). This raises an interesting tension between predetermined/confirmatory versus exploratory goals in descriptive research.

Still, because descriptive work tends to be considered exploratory rather than confirmatory, the value of preregistration is often questioned. Further, even if a goal can be specified in advance (e.g., comparing multimodal input to language learning), this analysis cannot be considered completely confirmatory if there is not a clear, predefined decision about which features of multimodal input will be annotated and how. This calls into question whether there is any value in preregistering a descriptive study at all. In reality, however, most work—descriptive included—falls on a continuum between confirmatory and exploratory (Scheel et al., 2021). It is the case that various dimensions of preregistration, like conducting a power analysis to determine sample size or outlining comprehensive analysis plans in advance, can be challenging for descriptive work. However, many dimensions of preregistration may still be valuable.

For example, even in descriptive work, researchers often have hypotheses and predictions about what they expect to find (see Kremin et al., 2020 or Romeo et al., 2018 for examples of preregistered studies that fall under our definition of descriptive research). Preregistration for descriptive work may include information about the behaviours the researchers seek to describe (e.g., if describing the occurrence of gestures, one can define what is meant by "gesture" and what counts or does not count as a gesture), how reliability will be assessed, what comparisons (if any) will be tested, and whether and how the video or audio corpus being described has been used in previous studies (which clarifies what researchers knew about the data before generating hypotheses). Additionally, researchers may opt to create simpler preregistrations (e.g., using the AsPredicted template or website; https://aspredicted.org/) that include answers to a few brief questions about study design and analyses (e.g., Kosie & Lew-Williams, in prep). Another option is to finalize a coding scheme with a few pilot participants and then—prior to coding the remaining data—preregister a study's methods (e.g., Kremin et al., 2020). If a researcher desires to preregister the initial coding as well, sequential preregistration (Nosek et al., 2018), in which components of a study can be preregistered incrementally, may be a useful option.

A common, yet incorrect, belief about preregistration is that decisions are set in stone and preclude changes or exploratory analyses (Nosek et al., 2019). This misconception can lead to researchers avoiding preregistration out of fear that they will want to change an aspect of their design or conduct additional analyses. It is often better to amend preregistration with justification for the amendment rather than to stick with, for example, a coding procedure or analysis that is not ideal. After all, the primary goal of preregistration is transparency, rather than imposing unnecessary constraints on research. Sharing this information enables readers of the paper and the preregistration to make judgements about both their confidence in different pieces of the work and also as a source of information about how the study was conducted and how to replicate it.

While we suggest that there are reasons why preregistration may be worthwhile for a descriptive study and encourage adoption of this practice, we acknowledge that the two are sometimes not a perfect fit. Researchers should make their own decisions about whether and what type of preregistration makes sense for their particular research design. The presence or absence of preregistration is not the sole determinant of whether or not a study is "open", and there are many other practices—a number of which are discussed below—that can be implemented by descriptive researchers to increase the transparency and replicability of their work. However, preregistration for a descriptive study can serve as a useful mechanism for outlining plans in advance and allows readers to decide for themselves where to put a study on the spectrum between exploratory and confirmatory (with neither end of the spectrum having objectively greater value).

5 | JUSTIFYING SAMPLE SIZE

Determining sample size can be challenging, even in experimental research when there is a strong expectation about the size and direction of an effect that a researcher expects to find. In descriptive work, there may be even weaker expectations about the size of an effect and, in some cases, the goal may be to simply describe variability in behaviour rather than look for a particular effect at all. Thus, a power analysis may not be possible or, if many parameters are unknown or require reliance on vague estimates, simply be uninformative. Further, even if enough information exists to enable a power analysis, it may not be possible to collect the sample size required, given that great effort is usually put into descriptions of even one participant. It can be challenging to recruit infants in general, and these challenges are more pronounced when the target of an investigation is an understudied population or a small cultural group. However, sample size issues cannot be completely ignored, and researchers should think carefully about the value of the information the study will provide for a given sample size (Lakens, 2021).

As with preregistration, the presence or absence of a power analysis is also not the sole determiner of the openness or transparency of research. When a power analysis is not possible for whatever reason, there are many other methods that can be used to justify the sample size (e.g., Lakens, 2021). For example, time may be a factor that limits a researcher to a relatively small sample. Manual transcription of audio or video files is a lengthy process (one estimate, for example, is that careful transcription takes 1 h per minute of speech; Gippert et al., 2006), and the time it takes to generate a full transcription can vary widely depending on features like the clarity of the audio or video files and the level of detail at which the files are being coded. It may also be the case that it is simply not feasible to collect data from more than a small number of individuals from an understudied population or culture (e.g., Casillas et al., 2020), or even more than just one individual in extreme cases (e.g., Roy et al., 2015). These limiting factors can be described as sample size justifications in a manuscript (and reported in advance in preregistration). Additionally, it can often be the case that having a small sample of data is preferable to having no data at all, as small samples can generate meaningful insights about behavioural phenomena—perhaps even more so in descriptive than in experimental studies. Transparency in how the sample size was determined, no matter how small or large enables readers to make an informed judgement about how to interpret descriptive research findings (Lakens, 2021).

As in experimental work (e.g., ManyBabies Consortium, 2020), descriptive researchers have embraced collaborative science to increase the sample size and cultural and linguistic diversity of research participants. For example, the Analysing Child Language Experiences around the World (ACLEW; Soderstrom et al., 2021) project is collaboration between researchers who build datasets, design computational tools and have developed a standardized transcription scheme for studying early language experiences in a way that is cross-culturally valid. While the ACLEW project primarily focuses on language and auditory input, the goal of the Play and Learning Across a Year (PLAY) project (Soska et al., 2021) is to build a dataset that includes videos of naturalistic activity between caregivers and infants. When complete, the PLAY dataset will contain coded videos of mother-infant interactions, video tours of families' homes, measures of ambient noise, and data from a variety of parent-report questionnaires collected by scientists at 50 universities in the United States and Canada. In addition to increasing sample size and diversity, researchers also benefit from the carefully constructed and shared coding procedures and tools for analysing descriptive data that are generated by collaborative efforts like these.

Beyond large-scale collaborations like ACLEW and PLAY, the sharing of data and materials for individual descriptive studies of any size (described in further detail below) can increase the cumulative sample size and diversity as well. For example, a single descriptive study may not have sufficient data to draw robust conclusions, and the sample characteristics may not allow generalization to other groups or cultures. However, if materials (e.g., videos, audio files, coding manuals, etc.) are shared, other researchers may be able to add to and build on this existing research. Cumulative science efforts like this enable researchers to amass large and diverse samples with which to investigate and explore variability surrounding a broad range of research questions. Each individual sample, no matter the size, adds incremental value.

6 | OPEN CODING PROCEDURES AND MATERIALS

Researchers doing descriptive work are interested in a large variety of behaviours, most of which do not have a one-size-fits-all definition. For example, caregivers' or infants' gestures have been of interest to many researchers (Gogate et al., 2000; Iverson et al., 1999; Kosie & Lew-Williams, in prep; Rowe et al., 2008), but the precise definition of a gesture may vary across studies. While some studies may count only pointing as a "gesture," others may include "any intentional, communicative movement of the body or limbs" (e.g., pointing *and* head nods *and* shrugging the shoulders for "I don't know" *and* extending and retracting the index finger to represent a worm; Iverson et al., 1999; Kosie & Lew-Williams, in prep). There are merits to both of these approaches, but what exactly counts as a gesture may differ across studies based on the researchers' goal, intuition or previous experience. To further complicate things, there may be variability across studies in basic considerations such as the timing at which a gesture is considered to have begun or ended. A comprehensive description of coding procedures is essential for reliability and replication, as well as a basic understanding and interpretation of research findings.

It quickly becomes apparent that clear and detailed coding instructions and guidelines matter. In the above example, the estimate of the number of gestures could vary widely depending on guidelines about when and how gestures are coded. A lack of clear coding guidelines can result in low reliability across coders within a given study or incorrect conclusions about how well the results of one study reporting the frequency of caregiver and infant gestures align with the results of another report of gesture frequency. As a consequence of this required degree of attention, it often takes extensive training for coders to become reliable in coding such observed behaviours. Within one lab, researchers often create detailed coding manuals that specify each consideration that a coder must make when deciding whether and when behaviour occurs. However, it is not uncommon for such descriptions to receive only a brief mention in a methods section. One contributing factor to this issue is that the level of detail that can be included in the methods section of papers is restricted by the word limit, often leading to a lack of clarity about exactly what counts as each type of coded behaviour. This makes it challenging for readers to understand precisely what was coded and to attempt to replicate the paper's coding in their own work. Additionally, as mentioned above, researchers can use and build on existing shared coding procedures. In addition to avoiding the need to "reinvent the wheel" each time a researcher wants to code a specific behaviour, use of the same shared coding schema (e.g., Soderstrom et al., 2021) makes it easier to compare how behaviour varies (or not) across studies or groups, and further increases the potential for later analyses that pool across datasets.

For most publications, coding materials can be shared in supplementary materials or posted online (e.g., on the Open Science Framework or OSF; https://www.osf.io/) for use by other researchers. Ideally, shared coding materials would include a detailed coding manual and example clips depicting the different coded behaviours. We have found that it can be helpful for training and clarity to have three types of clips: (1) examples that clearly depict the behaviour of interest, (2) examples that are clearly *not* the behaviour of interest, and (3) edge cases that challenge inter-coder reliability, along with a description of why each case does or does not count as the observed behaviour. These existing examples can be shared in supplementary material, after addressing participant confidentiality, to increase transparency in the description of coding procedures.

Beyond basic information about how coded variables were defined, transparency in timing-related information should also be included in studies that code video and/or audio data (Mendoza & Fausey, 2021b). For example, some researchers code continuously (by identifying the onset and offset of behaviours as a video unfolds; e.g., Kosie & Lew-Williams, in prep) while others ask coders to make binary decisions about events or objects present in shorter clips. For example, Fausey et al. (2016) asked coders to identify the presence of faces or hands in images extracted at ⅕ Hz (i.e., once every 5 s) from continuous video, while Cychosz et al. (2021) presented coders with 30-s audio clips which were coded for various features of infants' speech input. Still, others use automated analyses to identify regions of interest—such as moments with lots of speaking—and code only those regions (e.g., Weisleder & Fernald, 2013). To increase transparency and replicability, we and others (e.g., Mendoza & Fausey, 2021b) recommend including information about why a particular method was chosen along with evidence to support the choice (e.g., why would you expect to accurately capture a behaviour using 5 s clips versus continuous coding? Why 5 s and not 30 s clips?).

Technical details about coding can increase transparency as well. For example, knowing the software used to code video or audio files (e.g., Datavyu, Datavyu Team, 2014; ELAN, Wittenburg et al., 2006) can help others understand the coding structure, as well as details about opening and interpreting coded files. A number of researchers have shared even more detailed resources that include, for example, annotation procedures, accompanying materials and software and scripts. Mendoza and Fausey (2021b) outline in detail their procedures for annotating daylong audio recordings, including discussion of theoretical issues at stake as well as supporting materials for implementation (e.g., training manuals, codes for assessing reliability) available on an associated OSF page. This work is an ideal example of shared information about coding procedures and is an immensely helpful resource for researchers who want to annotate daylong audio recordings. The ACLEW team has also developed a variety of resources that include training materials for transcription as well as tools for automated annotation of daylong audio recordings (e.g., Le Franc et al., 2018; Räsänen et al., 2021; Soderstrom et al., 2021). Cychosz et al. (2021) created a set of Python scripts for efficiently sampling and annotating daylong audio recordings for various features of natural linguistic input (e.g., addressee, language); these scripts are described in detail in the manuscript and freely shared on GitHub (https://github.com/megseekosh/Categorize_app_v2). In addition to these examples, a variety of guidelines, scripts and packages have been created and shared by researchers to facilitate the collection, coding and analysis of naturalistic input (e.g., Anderson et al., 2021; Casillas & Scaff, 2021; Gautheron et al., 2021; Manning et al., 2020; Sanchez et al., 2019; Woon et al., 2021). All of these products serve as excellent examples of how to increase the openness and transparency of descriptive work, but they are also important resources that researchers can implement themselves when coding naturalistic behaviour.

Full transparency about coding also includes clear information about coder training and reliability, either in the manuscript itself or in supplementary materials (e.g., Mendoza & Fausey, 2021a). There is substantial variation in what type of reliability analysis is used and is appropriate across descriptive studies. Information related to reliability includes: how many files were coded by two or more coders, what counts as a “match” in coding, and what analyses were used to assess reliability. When coding gestures, does a code only count as “matching” if coders agree that the gesture starts and ends in the same second? For some types of descriptive work, the level of agreement we have come to expect from experimental tasks (e.g., looking to the left or right) may be unrealistic. For example, researchers may choose to loosen the criteria for a “match” between two coders (e.g., rather than requiring that a gesture be identified in exactly the same frame, a gesture counts as a match as long as both coders identify a gesture within a 2-s window). In addition to allowing readers to assess the reliability of coded behaviours, providing readers with details about how reliability was assessed is a valuable source of information about how to approach reliability when using the existing coding methods in their own work.

Finally, researchers' own positionality—their world views, experiences and the position they adopt toward a research study—has a substantial influence on their work, including the behaviours that they choose to code (Field & Derksen, 2021; Holmes, 2020; Malterud, 2001). Acknowledging this subjectivity can sometimes raise concerns or lead readers to devalue work that is more descriptive in nature, due to an outsize focus on objectivity as a hallmark of high-quality research (Field & Derksen, 2021; Gough & Madill, 2012). However, while subjectivity may be more obvious in descriptive studies, researchers' backgrounds and beliefs significantly bias experimental work as well (Gough & Madill, 2012). Additionally, the process of reflexivity—researchers' awareness of how personal experiences influence their research—can be extremely beneficial (Field & Derksen, 2021; Finlay, 2002a, 2002b). Acknowledging such influences can help the researcher understand and perhaps reduce their effects, and can sometimes lead to new insights that foster discovery. Making coding procedures and materials open can allow others to better assess how the original researchers' positionality may have influenced the conclusions of a study and perhaps adapt the coding in light of some of these effects. To make researchers' own individual influence(s) more explicit, they may choose to include positionality statements in their papers, as is common in much qualitative work (Holmes, 2020).

7 | SHARING AUDIO AND VIDEO FILES

Sharing audio and video files is common in descriptive work (e.g., Databrary., 2012; Gilmore et al., 2018; MacWhinney, 2000, 2014; MacWhinney & Snow, 1985; VanDam et al., 2016). In addition to simply making the research more transparent, sharing audio and video data has a number of benefits for the research community. For example, if one seeks to replicate the coding procedures of a previous study, it can be helpful to see video examples of moments at which the original researchers coded a particular behaviour of interest. Additionally, sharing audio and video files enables others to use these data to ask new questions or to extend existing research to a new culture or community. In this way, open audio and video data—along with procedural information such as instructions given to participants—can increase the accessibility of research, particularly for research groups who cannot easily collect this type of data. Sharing can also enable researchers to extend a method of coding that they created using their own audio or video data (perhaps collected from a convenience sample located near their own lab) to similar audio or video recordings of infants in different contexts or cultures, increasing the generalizability of findings.

Decisions about what to share may depend on the types of permissions allowed by the lab's Institutional Review Board (IRB) or ethics committee or on what each participant or community is willing to make publicly shareable. Under ideal circumstances, researchers would share (securely and with parental permission) unedited audio or video files along with coded data, as a reader could see first-hand how each behaviour was coded across a corpus. However, this is not always feasible due to a variety of considerations (e.g., understudied populations, IRB guidelines, etc), and participant consent and confidentiality are always the primary determinants of what can be shared. Even when sharing full audio and video files is not possible, however, there are still options for making data open in ways that are ethical and respectful of privacy concerns. For example, researchers may prefer to share clips of recordings that depict important features of the reported observations (e.g., examples of behaviours that were classified as gestures) from just a few participants who have provided consent to do so. Another option is to share only transcriptions or files with only coded data rather than raw video or audio files.

There are many repositories available for sharing audio and video recordings of infants' and children's everyday experiences, as well as associated data, metadata and/or analysis tools. A few popular options include the Child Language Data Exchange System (CHILDES), Homebank and Databrary. CHILDES was established in 1984 and was the first major effort to archive and share transcripts of children's language; in fact, it is a remarkably early example of an open-science effort. Since its creation, CHILDES has grown to include tens of thousands of publicly available transcripts of children's early language input and use in a number of languages and across typical and atypical development (https://childes.talkbank.org/; MacWhinney, 2000, 2014; MacWhinney & Snow, 1985). Further, Childes-db (Sanchez et al., 2019) is an open database and associated tools that were developed to increase accessibility and usability of the thousands of transcripts of natural parent–child interaction that are contained within the CHILDES database. Like CHILDES, Homebank is an online repository for recordings of child language with the focus of sharing daylong, naturalistic audio recordings along with metadata (e.g., family demographic information, transcriptions, output of automated speech processing programs) and tools—such as ACLEW-DARCLE Annotation Scheme and HomeBankCode—for annotating and analysing these large corpora (https://homebank.talkbank.org/; VanDam et al., 2016). Unlike CHILDES, however, Homebank implements strict user restrictions (e.g., files are stored in a password-protected format and accessible only to those who have agreed to confidentiality guidelines and have passed approved ethical training courses), as the daylong audio recordings it houses may be likely to contain personally identifying information about the families recorded. Finally, Databrary is a data library based at New York University that aims to "support data sharing among researchers in the behavioral, social, educational, developmental, neural and computer sciences" (https://nyu.databrary.org/; Databrary., 2012; Gilmore et al., 2018). Databrary stores audio and video files as well as many other types of experimental and demographic data, along with an accompanying tool for data coding and visualization (Datavyu, Datavyu Team, 2014). Access to files stored on Databrary can be restricted to any "Authorized Investigators" (e.g., researchers who have approval from their own institutions to access Databrary), shared only with the explicit approval of the data owner or kept completely private.

The ethical considerations when sharing data from descriptive studies are frequently more pronounced than when sharing data from experimental studies. While sharing full video or audio files best satisfies the goals of openness and transparency, this comes with a variety of issues surrounding participant confidentiality and privacy (see Cychosz et al., 2020; Gilmore et al., 2020, 2021). Video and audio files frequently contain identifiable information like faces or names that cannot be as easily omitted as they can in experimental data files. Further, video and audio recordings collected in participants' homes and the surrounding community may raise additional concerns about protecting the privacy of not only the participants who have consented to be part of the study but also individuals they may naturally encounter during the recording but who may not have consented to be recorded. These issues have received extensive consideration by organizations that facilitate the sharing of video and audio recordings. For example, Databrary (Databrary., 2012; Gilmore et al., 2018) provides templates and guidelines for obtaining parental informed consent. Both Databrary and Homebank (VanDam et al., 2016) offer various levels of data sharing, enabling researchers to decide whether to make data publicly available or only available to researchers who have specifically requested access to a particular corpus.

Two additional concerns may arise concerning *when* to share data. From a practical perspective, it can feel overwhelming to think about how to find the time to prepare data for sharing along with all of the other tasks that come with collecting a corpus and preparing a paper for submission. However, the time costs associated with making audio and video data open can be minimized if researchers start from the beginning with the intention to share. For example, videos and associated metadata can be immediately uploaded to a private Databrary or Homebank repository once parental consent is received. By the time data from all sessions are collected or the researcher is ready to make the data open, the repository will be ready for sharing. Additionally, collecting and annotating a corpus is an enormous endeavour that requires a great deal of time, technical and intellectual effort, and it is still not clear how to appropriately credit those who do the extensive work of creating a large corpus (e.g., Gennetian et al., 2022). Thus, it may be desirable to embargo data (i.e., wait to share until the creators of the corpus have fully met their research goals). Alternatively, researchers can limit what they make available (e.g., audio but not video files or annotations alone) until they are ready to share the entire corpus or share the corpus only with those who have contacted them directly to request access. While it may be optimal to share data from the start, these alternatives allow researchers to openly share data in the way that works best for them.

8 | FINAL CONSIDERATIONS

Throughout this paper, we have discussed various practices that increase the openness and transparency of descriptive research, and, in conclusion, we articulate a number of final considerations: (1) Implementing all of these guidelines at once can be overwhelming, and it is likely that our list does not even include all possible practices for increasing openness and transparency (for example, while not typically considered a component of open science, detailed information about the characteristics of a sample both increases transparency and reduces inappropriate generalization of findings). However, it is important to acknowledge that engaging in even one practice that increases transparency is a step in the right direction, and additional practices can be incrementally added over time (Bergmann, 2018; Corker, 2018; Kathawalla et al., 2021; Nuijten, 2019; Syed, 2019). (2) It should be acknowledged that these practices increase transparency more broadly, not just in descriptive work. For example, sharing videos also increases transparency and the ability to replicate experimental studies, and the necessity of introspection and analysis of researchers' own positionality extends to all domains of research. (3) Our suggestions for open-science practices in descriptive research were informed by and adapted from practices that were predominantly targeted toward experimental studies. A useful future endeavour would be to essentially "start from scratch," ignoring all previous open-science recommendations and generating an entirely new conceptualization of what open science means in descriptive research. (4) Finally, we encourage readers to refrain from making assumptions about the quality of work based only on the presence or absence of open-science practices (Simmons et al., 2021). Papers that check all of the boxes for open science can still be problematic, and many wonderful papers exist that implement few or no open-science practices. Regardless of open-science practices, readers should evaluate all available information when assessing the quality of scientific research.

9 | CONCLUSION

Descriptive research is important for a variety of reasons: describing behaviours and phenomena, constructing hypotheses, generating ideas for experiments, informing computational models of input and output, and answering questions about differences and similarities across groups, communities or cultures. Doing descriptive research in a way that is open and transparent promotes reproducibility, enhances readers' ability to interpret results, and sets up other researchers to answer similar research questions in ways that are consistent across labs (e.g., using existing coding methods and/or existing data). In this paper, we discuss four practices for increasing transparency: preregistration, justifying sample size, making coding procedures open, and sharing video and audio data. For each practice, we describe the benefits of openness and transparency and make suggestions for how to implement open-science practices. Throughout, we discuss various ongoing challenges that we hope will prompt thoughtful practices by researchers who want to do open, transparent descriptive research. In summary, descriptive research is vital to developmental science. Openness and transparency improve the reliability and replicability of descriptive work, thus increasing confidence in research findings. Perhaps more importantly, though, is that the practice of sharing procedures, data and tools allows researchers to learn from one another and work together to advance fundamental understanding of early learning and development.

AUTHOR CONTRIBUTIONS

Jessica E. Kosie: Conceptualization; writing – original draft; writing – review and editing. **Casey Lew-Williams:** Conceptualization; supervision; writing – review and editing.

CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1002/icd.2377.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed.

ORCID

Jessica E. Kosie https://orcid.org/0000-0002-2390-0963

REFERENCES

Adolph, K. E., Hoch, J. E., & Cole, W. G. (2018). Development (of walking): 15 suggestions. *Trends in Cognitive Sciences*, *22*(8), 699–711.

Adolph, K. E., Cole, W. G., Komati, M., Garciaguirre, J. S., Badaly, D., Lingeman, J. M., Chan, G. L., & Sotsky, R. B. (2012). How do you learn to walk? Thousands of steps and dozens of falls per day. *Psychological Science*, *23*(11), 1387–1394. https://doi.org/10.1177/0956797612446346

Adolph, K. E., & Tamis-LeMonda, C. S. (2014). The costs and benefits of development: The transition from crawling to walking. *Child Development Perspectives*, *8*(4), 187–192. https://doi.org/10.1111/cdep.12085

Anderson, H.T., Mendoza, J.K., Imhof, A., & Fausey, C.M. (2021). ITSbin: Wrangle LENA ITS files into linear seconds & minutes. R package version 1.0.0. https://doi.org/10.5281/zenodo.4474553

Aslin, R. N. (2007). What's in a look? *Developmental Science*, *10*(1), 48–53. https://doi.org/10.1111/j.1467-7687.2007.00563.x

Bergelson, E., Amatuni, A., Dailey, S., Koorathota, S., & Tor, S. (2019). Day by day, hour by hour: Naturalistic language input to infants. *Developmental Science*, *22*(1), e12715. https://doi.org/10.1111/desc.12715

Bergelson, E., Casillas, M., Soderstrom, M., Seidl, A., Warlaumont, A. S., & Amatuni, A. (2019). What do north American babies hear? A large-scale cross-corpus analysis. *Developmental Science*, *22*(1), e12724. https://doi.org/10.1111/desc.12724

Bergmann, C. (2018). How to integrate open science into language acquisition research. Student workshop at the 43rd Boston University conference on language development (BUCLD).

Bergmann, C., Tsuji, S., Piccinini, P. E., Lewis, M. L., Braginsky, M., Frank, M. C., & Cristia, A. (2018). Promoting replicability in developmental research through meta-analyses: Insights from language acquisition research. *Child Development*, *89*(6), 1996–2009. https://doi.org/10.1111/cdev.13079

Braun, V., & Clarke, V. (2021). To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. *Qualitative Research in Sport, Exercise and Health*, *13*(2), 201–216. https://doi.org/10.1080/2159676X.2019.1704846

Bunce, J., Soderstrom, M., Bergelson, E., Rosemberg, C., Stein, A., Migdalek, M., & Casillas, M. (2020). A cross-cultural examination of young children's everyday language experiences. PsyArXiv. https://doi.org/10.31234/osf.io/723pr

Byers-Heinlein, K., Bergmann, C., Davies, C., Frank, M. C., Hamlin, J. K., Kline, M., Kominsky, J. F., Kosie, J. E., Lew-Williams, C., Liu, L., Mastroberardino, M., & Soderstrom, M. (2020). Building a collaborative psychological science: Lessons learned from ManyBabies 1. *Canadian Psychology/Psychologie Canadienne*, *61*(4), 349–363. https://doi.org/10.1037/cap0000216

Byers-Heinlein, K., Bergmann, C., & Savalei, V. (2021). Six solutions for more reliable infant research. *Infant and Child Development*, e2296. https://doi.org/10.1002/icd.2296

Casey, K., Potter, C., Lew-Williams, C., & Wojcik, E. (in prep). Moving beyond "nouns in the lab": Using naturalistic data to understand why infants' first words include uh-oh and hi.

Casillas, M., Brown, P., & Levinson, S. C. (2020). Early language experience in a Tseltal Mayan village. *Child Development*, *91*(5), 1819–1835. https://doi.org/10.1111/cdev.13349

Casillas, M., & Scaff, C. (2021). Analyzing contingent interactions in R with chattr. In proceedings of the annual meeting of the cognitive science society (Vol. 43, No. 43).

Clerkin, E. M., Hart, E., Rehg, J. M., Yu, C., & Smith, L. B. (2017). Real-world visual statistics and infants' first-learned object names. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711), 20160055. https://doi.org/10.1098/rstb.2016.0055

Corker, K. S. (2018). Open Science is a Behavior. Center of Open Science. https://cos.io/blog/open-science-is-a-behavior

Custode, S. A., & Tamis-LeMonda, C. (2020). Cracking the code: Social and contextual cues to language input in the home environment. *Infancy*, *25*(6), 809–826. https://doi.org/10.1111/infa.12361

Cychosz, M., Romeo, R., Soderstrom, M., Scaff, C., Ganek, H., Cristia, A., Casillas, M., De Barbaro, K., Bang, J. Y., & Weisleder, A. (2020). Longform recordings of everyday life: Ethics for best practices. *Behavior Research Methods*, *52*(5), 1951–1969. https://doi.org/10.3758/s13428-020-01365-9

Cychosz, M., Villanueva, A., & Weisleder, A. (2021). Efficient estimation of children's language exposure in two bilingual communities. *Journal of Speech, Language, and Hearing Research*, *64*(10), 3843–3866. https://doi.org/10.1044/2021_JSLHR-20-00755

Darwin, C. (1877). A biographical sketch of an infant. *Mind*, *2*(7), 285–294.

Databrary. (2012). The databrary project: A video data library for developmental science [computer software manual]. http://www.databrary.org

Datavyu Team. (2014). *Datavyu: A video coding tool. Databrary project*. New York University. datavyu.org

Davis-Kean, P. E., & Ellis, A. (2019). An overview of issues in infant and developmental research for the creation of robust and replicable science. *Infant Behavior and Development*, *57*, 101339. https://doi.org/10.1016/j.infbeh.2019.101339

DeBolt, M. C., Rhemtulla, M., & Oakes, L. M. (2020). Robust data and power in infant research: A case study of the effect of number of infants and number of trials in visual preference procedures. *Infancy*, *25*(4), 393–419. https://doi.org/10.1111/infa.12337

Eason, A. E., Hamlin, J. K., & Sommerville, J. A. (2017). A survey of common practices in infancy research: Description of policies, consistency across and within labs, and suggestions for improvements. *Infancy*, *22*(4), 470–491. https://doi.org/10.1111/infa.12183

Fausey, C. M., Jayaraman, S., & Smith, L. B. (2016). From faces to hands: Changing visual input in the first two years. *Cognition*, *152*, 101–107. https://doi.org/10.1016/j.cognition.2016.03.005

Field, S. M., & Derksen, M. (2021). Experimenter as automaton; experimenter as human: Exploring the position of the researcher in scientific research. *European Journal for Philosophy of Science*, *11*(1), 1–21. https://doi.org/10.1007/s13194-020-00324-7

Finlay, L. (2002a). "Outing" the researcher: The provenance, process, and practice of reflexivity. *Qualitative Health Research*, *12*(4), 531–545.

Finlay, L. (2002b). Negotiating the swamp: The opportunity and challenge of reflexivity in research practice. *Qualitative Research*, *2*(2), 209–230.

Frank, M. C., Bergelson, E., Bergmann, C., Cristia, A., Floccia, C., Gervain, J., Hamlin, J. K., Hannon, E. E., Kline, M., Levelt, C., Lew-Williams, C., & Yurovsky, D. (2017). A collaborative approach to infant research: Promoting reproducibility, best practices, and theory-building. *Infancy*, *22*(4), 421–435. https://doi.org/10.1111/infa.12182

Gautheron, L., Rochat, N., & Cristia, A. (2021). Managing, storing, and sharing long-form recordings and their annotations. PsyArXiv. https://doi.org/10.31234/osf.io/w8trm

Gennetian, L. A., Frank, M., & Tamis-LeMonda, C. S. (2022). Open Science in developmental science. *Annual Review of Developmental Psychology*, *4*.

Gesell, A. (1928). *Infancy and human growth*. Macmillan.

Gesell, A. (1940). *The first five years of life*. Harper.

Gesell, A., & Ilg, F. L. (1943). *Infant and child in the culture of today*. Harper.

Gesell, A., & Thompson, H. (1934). *Infant behavior: Its genesis and growth*. McGraw-Hill.

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

Gilmore, R. O., Cole, P. M., Verma, S., Van Aken, M. A., & Worthman, C. M. (2020). Advancing scientific integrity, transparency, and openness in child development research: Challenges and possible solutions. *Child Development Perspectives*, *14*(1), 9–14. https://doi.org/10.1111/cdep.12360

Gilmore, R. O., Xu, M., & Adolph, K. E. (2021). Data sharing. In S. Panicker & B. Stanley (Eds.), *How to conduct research ethically*. American Psychological Association.

Gippert, J., Himmelmann, N., & Mosel, U. (2006). *Essentials of language documentation*. Mouton de Gruyter. https://doi.org/10.1515/9783110197730

Gogate, L. J., Bahrick, L. E., & Watson, J. D. (2000). A study of multimodal motherese: The role of temporal synchrony between verbal labels and gestures. *Child Development*, *71*(4), 878–894. https://doi.org/10.1111/1467-8624.00197

Gough, B., & Madill, A. (2012). Subjectivity in psychological science: From problem to prospect. *Psychological Methods*, *17*(3), 374–384. https://doi.org/10.1037/a0029313

Haven, T. L., Errington, T. M., Gleditsch, K. S., van Grootel, L., Jacobs, A. M., Kern, F. G., Piñeiro, R., Rosenblatt, F., & Mokkink, L. B. (2020). Preregistering qualitative research: A Delphi study. *International Journal of Qualitative Methods*, *19*, 1609406920976417. https://doi.org/10.1177/1609406920976417

Havron, N., Bergmann, C., & Tsuji, S. (2020). Preregistration in infant research—A primer. *Infancy*, *25*(5), 734–754. https://doi.org/10.1111/infa.12353

Herzberg, O., Fletcher, K. K., Schatz, J. L., Adolph, K. E., & Tamis-LeMonda, C. S. (2021). Infant exuberant object play at home: Immense amounts of time-distributed, variable practice. *Child Development.*, *93*, 150–164. https://doi.org/10.1111/cdev.13669

Hoch, J. E., O'Grady, S. M., & Adolph, K. E. (2019). It's the journey, not the destination: Locomotor exploration in infants. *Developmental Science*, *22*(2), e12740. https://doi.org/10.1111/desc.12740

Hoch, J. E., Rachwani, J., & Adolph, K. E. (2020). Where infants go: Real-time dynamics of locomotor exploration in crawling and walking infants. *Child Development*, *91*(3), 1001–1020. https://doi.org/10.1111/cdev.13250

Holmes, A. G. D. (2020). Researcher Positionality--A Consideration of Its Influence and Place in Qualitative Research--A New Researcher Guide. *Shanlax International Journal of Education*, *8*(4), 1–10. https://doi.org/10.34293/education.v8i4.3232

Humphreys, L., Lewis, N. A., Sender, K., & Won, A. S. (2021). Integrating qualitative methods and open science: Five principles for more trustworthy research. *Journal of Communication.*, *71*(5), 855–874. https://doi.org/10.1093/joc/jqab026

Iverson, J. M., Capirci, O., Longobardi, E., & Caselli, M. C. (1999). Gesturing in mother-child interactions. *Cognitive Development*, *14*(1), 57–75. https://doi.org/10.1016/S0885-2014(99)80018-5

Karasik, L. B., Tamis-LeMonda, C. S., & Adolph, K. E. (2011). Transition from crawling to walking and infants' actions with objects and people. *Child Development*, *82*(4), 1199–1209. https://doi.org/10.1111/j.1467-8624.2011.01595.x

Karasik, L. B., Tamis-LeMonda, C. S., Ossmy, O., & Adolph, K. E. (2018). The ties that bind: Cradling in Tajikistan. *PLoS One*, *13*(10), e0204428. https://doi.org/10.1371/journal.pone.0204428

Karasik, L. B., Tamis-LeMonda, C. S., & Adolph, K. E. (2014). Crawling and walking infants elicit different verbal responses from mothers. *Developmental Science*, *17*(3), 388–395. https://doi.org/10.1111/desc.12129

Kathawalla, U. K., Silverstein, P., & Syed, M. (2021). Easing into open science: A guide for graduate students and their advisors. *Collabra: Psychology*, *7*(1). https://doi.org/10.1525/collabra.18684

Kosie, J. E., & Lew-Williams, C. (in prep). Infant-Directed Communication: Speech, Action, Gesture, Emotion, and Touch in Caregiver-Infant Interactions.

Kremin, L. V., Alves, J., Orena, A. J., Polka, L., & Byers-Heinlein, K. (2020). Code-switching in parents' everyday speech to bilingual infants. *Journal of Child Language*, *1-27*, 714–740. https://doi.org/10.1017/S0305000921000118

Lakens, D. (2021). Sample size justification. PsyArXiv. https://doi.org/10.31234/osf.io/9d3yf

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 INTERSPEECH 1383-1387.

MacWhinney, B. (2000). *The CHILDES project: The database* (Vol. 2). Psychology Press.

MacWhinney, B. (2014). *The CHILDES project: Tools for analyzing talk*, Volume II. The database. Psychology Press.

MacWhinney, B., & Snow, C. (1985). The child data exchange system. *Journal of Child Language*, *12*(2), 271–295. https://doi.org/10.1017/S0305000900006449

Malterud, K. (2001). Qualitative research: Standards, challenges, and guidelines. *The Lancet*, *358*(9280), 483–488. https://doi.org/10.1016/S0140-6736(01)05627-6

Manning, B. L., Harpole, A., Harriott, E. M., Postolowicz, K., & Norton, E. S. (2020). Taking language samples home: Feasibility, reliability, and validity of child language samples conducted remotely with video chat versus in-person. *Journal of Speech, Language, and Hearing Research*, *63*(12), 3982–3990. https://doi.org/10.1044/2020_JSLHR-20-00202

ManyBabies Consortium. (2020). Quantifying sources of variability in infancy research using the infant-directed-speech preference. *Advances in Methods and Practices in Psychological Science*, *3*(1), 24–52. https://doi.org/10.1177/2515245919900809

McGraw, M. B. (1935). *Growth: A study of Johnny and Jimmy (preface by F. Tilney; introduction by J. Dewey.)*. Appleton-Century.

Mendoza, J. K., & Fausey, C. M. (2021a). Everyday music in infancy. *Developmental Science*, *24*(6), e13122. https://doi.org/10.1111/desc.13122

Mendoza, J. K., & Fausey, C. M. (2021b). Quantifying everyday ecologies: Principles for manual annotation of many hours of Infants' lives. *Frontiers in Psychology*, *12*, 710636. https://doi.org/10.3389/fpsyg.2021.710636

Nelson, L. D., Simmons, J., & Simonsohn, U. (2018). Psychology's renaissance. *Annual Review of Psychology*, *69*, 511–534. https://doi.org/10.1146/annurev-psych-122216-011836

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, *115*(11), 2600–2606. https://doi.org/10.1073/pnas.1708274114

Nosek, B. A., Beck, E. D., Campbell, L., Flake, J. K., Hardwicke, T. E., Mellor, D. T., van't Veer, A. E., & Vazire, S. (2019). Preregistration is hard, and worthwhile. *Trends in Cognitive Sciences*, *23*(10), 815–818. https://doi.org/10.1016/j.tics.2019.07.009

Nuijten, M. B. (2019). Practical tools and strategies for researchers to increase replicability. *Developmental Medicine & Child Neurology*, *61*(5), 535–539.

Oakes, L. M. (2017). Sample size, statistical power, and false conclusions in infant looking-time research. *Infancy*, *22*(4), 436–469. https://doi.org/10.1111/infa.12186

Peterson, D. (2016). The baby factory: Difficult research objects, disciplinary standards, and the production of statistical significance. *Socius*, *2*, 1–10. https://doi.org/10.1177/2378023115625071

Prather, R. W. (2021). Fear not of cognition in context. *Infant and Child Development*, *31*, e2249. https://doi.org/10.1002/icd.2249

Preyer, W. T. (1888). *The mind of the child: Observations concerning the mental development of the human being in the first years of life* (Vol. 7). D. Appleton and Company.

Räsänen, O., Seshadri, S., Lavechin, M., Cristia, A., & Casillas, M. (2021). ALICE: An open-source tool for automatic measurement of phoneme, syllable, and word counts from child-centered daylong recordings. *Behavior Research Methods*, *53*(2), 818–835. https://doi.org/10.3758/s13428-020-01460-x

Romeo, R. R., Segaran, J., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Yendiki, A., Rowe, M. L., & Gabrieli, J. D. (2018). Language exposure relates to structural neural connectivity in childhood. *Journal of Neuroscience*, *38*(36), 7870–7877. https://doi.org/10.1523/JNEUROSCI.0484-18.2018

Rowe, M. L., Özçalışkan, Ş., & Goldin-Meadow, S. (2008). Learning words by hand: Gesture's role in predicting vocabulary development. *First Language*, *28*(2), 182–199. https://doi.org/10.1177/0142723707088310

Roy, B. C., Frank, M. C., DeCamp, P., Miller, M., & Roy, D. (2015). Predicting the birth of a spoken word. *Proceedings of the National Academy of Sciences*, *112*(41), 12663–12668. https://doi.org/10.1073/pnas.1419773112

Sanchez, A., Meylan, S. C., Braginsky, M., MacDonald, K. E., Yurovsky, D., & Frank, M. C. (2019). Childes-db: A flexible and reproducible interface to the child language data exchange system. *Behavior Research Methods*, *51*(4), 1928–1941. https://doi.org/10.3758/s13428-018-1176-7

Scheel, A. M., Tiokhin, L., Isager, P. M., & Lakens, D. (2021). Why hypothesis testers should spend less time testing hypotheses. *Perspectives on Psychological Science*, *16*(4), 744–755. https://doi.org/10.1177/1745691620966795

Schott, E., Rhemtulla, M., & Byers-Heinlein, K. (2019). Should I test more babies? Solutions for transparent data peeking. *Infant Behavior and Development*, *54*, 166–176. https://doi.org/10.1016/j.infbeh.2018.09.010

Shinn, M. W. (1900). *The biography of a baby*. Houghton, Mifflin.
Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366. https://doi.org/10.1177/0956797611417632
Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2021). Pre-registration: Why and how. *Journal of Consumer Psychology*, *31*(1), 151–162. https://doi.org/10.1002/jcpy.1208
Soderstrom, M., Casillas, M., Bergelson, E., Rosemberg, C., Alam, F., Warlaumont, A. S., & Bunce, J. (2021). Developing a cross-cultural annotation system and MetaCorpus for studying Infants' real world language experience. *Collabra: Psychology*, *7*(1), 23445. https://doi.org/10.1525/collabra.23445
Soska, K., Xu, M., Gonzalez, S., Hertzberg, O., Gilmore, R. O., Tamis-LeMonda, C., & Adolph, K. E. (2021). (hyper) active data curation: A video case study from behavioral science. PsyArXiv. https://doi.org/10.31234/osf.io/89rcb
Spellman, B. A., Gilbert, E. A., & Corker, K. S. (2018). Open science. *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience*, *5*, 1–47. https://doi.org/10.1002/9781119170174.epcn519
Sullivan, J., Mei, M., Perfors, A., Wojcik, E., & Frank, M. C. (2020). SAYCam: A large, longitudinal audiovisual dataset recorded from the infant's perspective. *Open Mind*, *1-10*, 20–29. https://doi.org/10.1162/opmi_a_00039
Syed, M. (2019). The open science movement is for all of us. PsyArXiv. https://doi.org/10.31234/osf.io/cteyb
Thelen, E., & Adolph, K. E. (1992). Arnold L. Gesell: The paradox of nature and nurture. *Developmental Psychology*, *28*(3), 368–380.
Tincoff, R., Seidl, A., Buckley, L., Wojcik, C., & Cristia, A. (2019). Feeling the way to words: Parents' speech and touch cues highlight word-to-world mappings of body parts. *Language Learning and Development*, *15*(2), 103–125. https://doi.org/10.1080/15475441.2018.1533472
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. Thieme Medical Publishers. https://doi.org/10.1055/s-0036-1580745
Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014). A social feedback loop for speech development and its reduction in autism. *Psychological Science*, *25*(7), 1314–1324. https://doi.org/10.1177/0956797614531023
Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, *24*(11), 2143–2152. https://doi.org/10.1177/0956797613488145
Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: A professional framework for multimodality research. In 5th international conference on language resources and evaluation (LREC 2006)1556-1559.
Woon, F. T., Yogarrajah, E. C., Fong, S., Salleh, N. S. M., Sundaray, S., & Styles, S. J. (2021). Creating a corpus of multilingual parent-child speech remotely: Lessons learned in a large-scale onscreen picturebook sharing task. PsyArXiv. https://doi.org/10.31234/osf.io/yeg6u

How to cite this article: Kosie, J. E., & Lew-Williams, C. (2024). Open science considerations for descriptive research in developmental science. *Infant and Child Development*, *33*(1), e2377. https://doi.org/10.1002/icd.2377