A Report on the New Jersey Families Study's Focus Groups from December 2020

Cecilia H. Kim
New Jersey Families Study
Office of Population Research
Princeton University

January 2021

The New Jersey Families Study is located in the Office of Population Research, Wallace Hall, Princeton University, Princeton NJ 08544. Funding for the NJFS focus groups and for the preparation of this report was provided by the Data Driven Social Science Initiative at Princeton University. I am grateful to focus group participants and members of the NJFS Princeton team for helpful comments on an earlier version of this report.

A Report on the New Jersey Families Study's Focus Groups from December 2020

INTRODUCTION

The New Jersey Families Study (NJFS) is a video ethnographic examination of how families support their children's early learning. Our data archive holds a set of two-week, in-home video recordings of the daily lives of 21 highly diverse, young families in Mercer County, New Jersey. Altogether we have collected an impressive amount of data—roughly 5,700 hours of video recordings contained in approximately 460,000 discrete video clips—that focus on routine interactions between adult caregivers and 2- and 3-year-old children. Our goal is to create from this information a "user-friendly" early childhood database and make it available to a worldwide research community.[1]

To advance this goal, the project convened two focus groups in early December 2020 with ten early childhood experts. Each focus group had five participants and was organized around three topics. First, and most important, participants were asked for their views on the most useful pre-analytic behavioral content codes or descriptors of what is happening in the videos. Suppose that a researcher is interested in episodes of children crying. It would take at least eight months of round-the-clock viewing to go through all 460,000 clips and tag those where the target child is crying. No one is likely to invest all that time, and therein lies the problem. We need a faster way for researchers to identify the subset of all video clips that meets their requirements.

Our plan is to pre-code the video clips using a manageable set of broad content categories that are likely to capture a wide range of early childhood researchers' needs and then store the pre-coded clips in a searchable database. Using keywords or phrases—for example, "the target child is crying"—researchers can then almost instantaneously call up just those clips of special interest and proceed to do their own much more granular coding. The broad categories we choose are then just convenient off-ramps to a useful subset of the data.

A second topic centered around satisfactory levels of resolution for the video clips. Higher levels of resolution produce greater video quality, but they also increase storage requirements and therefore project costs. Finally, participants were asked for their wish list of other desirable characteristics of a user-friendly early childhood data set.

The remainder of this report summarizes the discussion in both focus groups. It is organized into an Executive Summary, a fuller accounting of participants' deliberations, and a set of appendices.

---

[1] More information about the study is contained in NJFS newsletters (https://scholar.princeton.edu/tje/newsletter).

EXECUTIVE SUMMARY

The New Jersey Families Study (NJFS) invited ten early childhood experts to participate in two focus groups in December 2020 in order to receive input on the development of a coding structure for the project's video data set and other insights about how the data could be best organized to support work of those in the field. During the focus groups, each comprised of five experts, participants were first asked to suggest macrocodes that could be used to create a topical index for pre-analytic coding of the NJFS data. The proposed codes were grouped under three broad headings: "Participants," "Activities and Behaviors," and "Interaction Bouts." In addition to the codes proposed by the outside experts, the NJFS Princeton team suggested metadata that provide basic information about the video clips. A full list of proposed metadata codes can be found in Appendix C.

The "Participants" classification pertains to information about the individuals who are present in the clips as well as their relationship to the target child. Focus group members agreed that these codes would facilitate studies that examine the effect of interaction with others, or the lack thereof, on early childhood development. If the "Participants" classification can be described as the "who" of the data, the "Activities and Behaviors" classification would be the "what." This category is more extensive and perhaps more complex, because it contains an element of subjectivity. This report describes the activity and behavioral codes that were frequently proposed by members of both groups, including as examples media and screen time, mealtime, playtime and cognitive stimulation, sleeping, speech, and emotion. The "Interaction Bouts" category combines the previous two classifications to examine when a child is interacting with a given person in an activity.

Following the discussion on codes for the topical index, focus group members were shown the same video clip (that only contained members of the NJFS setup team and no participating family members in compliance with NJFS privacy promises) at low, medium, and high resolutions to receive feedback on which resolution would be optimal for research and analysis, given the considerations of increasing cost and storage with increasing resolution. Because focus group members found it difficult to detect major differences due to the limitations of video sharing on the Zoom platform, a Dropbox link containing the videos was shared afterwards. There was a general consensus that the medium or high resolution would offer considerable advantages for viewing and analysis.

Lastly, participants were asked if they had any suggestions about other features of a user-friendly database that could potentially be applied to the NJFS data set. There were various suggestions, such as a desire for transcripts of the audio, simultaneous split-screen viewing for clips that were recorded at the same time in different parts of the house, a dynamic and searchable archive of microcodes that other researchers applied to the data, the ability to link videos that were connected by content or references, a measure of proximity between people, the ability to watch clips adjacent in time to study transitions, and the application of artificial intelligence to data preparation. Detailed findings from the discussions of the two focus groups, as well as appendices including information on the participants and metadata codes, follow.

DETAILED FINDINGS

Focus Group Methodology

The NJFS sent email invitations to ten early childhood experts, representing various specialties in their fields and different institutions. All ten agreed to participate; honoraria were offered to the participants. Two focus groups, each with five experts, were held over Zoom in December 2020, with the dates and times of the meetings as well as the names, affiliations, and brief biographies of participants included in Appendices A and B. Members of the NJFS Princeton team were also present at both focus groups as director, moderator, video presenter, rapporteur, and IT support. The meetings were recorded.

Suggested Codes for a Topical Index

Participants

This first category of codes would indicate which individuals are present in the video clip and their relationship to the target child, such as "target child," "mother," "father," "grandparent," "sibling," and "non-family member." This coding system would allow researchers to select for time periods when the target child is alone, when there are only adults present, when there are other people in the target child's presence, etc. and create their own microcodes. This could be a challenging programming task because who's there in the clip has a dynamic property. People can be entering and leaving a given room during the same video clip.

Activity and Behavior

If the previous coding classification can be described as the "who" portion of the video clips, this category would be the "what." Although behavioral coding is challenging due to its subjective nature (for example, should "eating" include only organized mealtimes or be expanded to include a child picking up a Cheerio from off the floor), the activities that were mentioned in both focus groups are outlined below. While the target child's perspective is primary here, there would likely be an appreciable benefit to coding from adult perspectives as well.

Media and Screen Time

A code that was proposed by several participants was exposure to and presence of media in a variety of forms (television, phone, computer, etc.), even in instances where the target child was not directly interacting with the media, for example, if the television was left on throughout the day.

Mealtime

This code would tag organized mealtimes, including breakfast, lunch, dinner, snack times between meals, and generalized eating. Mealtime codes would be helpful as

they would also likely correlate with periods of high adult-child interactions. The time of day associated with different meal times was also suggested as an organic way to decide the timescale of videos if the discrete clips were "stitched together" into longer segments.

Grooming

Participants also suggested coding for self-grooming, such as when a child is getting dressed, getting ready to go out, or getting ready for bed.

Movement

Movement would include a target child's wandering around, walking, running, crawling, and playful action.

Playtime and Cognitive Stimulation

In addition to coding for times when the child was engaged in play that was either guided by another individual or carried out independently, it would be beneficial to identify objects that are commonly interacted with during playtime such as pets, games, dolls, and toys. These could be tagged to study the development of skills such as object manipulation. The "playtime" code was grouped with the "cognitive stimulation" code so that researchers would be able to identify the variety of activities that are associated with both play and cognitive development such as solving puzzles, singing songs, or being read to, and make their own decisions as to if and how they wanted to distinguish these activities.

Sleeping

Coding for when the target child is sleeping, waking up, or going to bed would allow researchers to obtain information such as wake-up and bedtime schedules and routines. This code would also include instances of napping. In households where families gave consent, infrared cameras were installed in the target child's bedroom so that researchers would be able to observe motion during the night.

Speech

This code would indicate when there is speech or talking in the video clips in order to observe interactions among individuals and a child's language development. Having clear audio is important, especially because it is generally difficult to discern what young children are saying. One could also code for non-English speech.

Physical Touch

This code would tag physical interaction between the child and another individual, including social touch, being cuddled, moments of affirmation through physical touch, as well as instances when the child is subject to physical punishment.

Emotions and Emotive Action

Basic emotions such as happiness, sadness, anger, excitement, and fear could be tagged using aspects of the videos such as facial expression and speech. Participants also suggested coding for displays of emotive action such as yelling, kissing, crying, screaming, and throwing or banging an object, as well as developing a rating scale for valence and arousal.

Interaction Bouts

This category combines the previous two classifications to examine when a child is interacting with a given person in an activity. This coding system would allow researchers to create the necessary microcodes to analyze the target child's behavior during joint engagement, serve and return interactions, moments of potential connection and reciprocity, as well as the kinds of support for self-regulation, executive function, and related behaviors and skills. Rather than identifying these episodes directly, it may be possible to derive an approximate classification from separate filters for the activities and participants by applying them both simultaneously.

Feedback on Video Clip Resolution

After the topical index discussion, focus groups were shown a video clip (that only contained members of the NJFS setup team and no participating family members in compliance with NJFS privacy promises) at three different resolutions--480p (low), 720p (medium), and 1080p (high). Feedback from the focus members on which of the three resolutions would be most suitable for research purposes was valuable because a decision has implications for video storage space and cost. Given the quality limitations imposed by Zoom, focus group members found it difficult to detect large differences among the three resolutions. Subsequently, a Dropbox link containing the original videos was shared with the participants to receive their input. There were multiple comments that the video was "pixelated" at the lowest resolution, which could hinder research surrounding "micro-expressions and gestures as well as certain aspects of body coordination in play." There was a general consensus that the medium or high resolution would offer considerable advantages for viewing and analysis. Several participants observed that there was a noticeable difference between the low and medium resolutions, but not as much distinction between the medium and high resolutions.

Other Desirable Features of a User-Friendly Data Set

Participants were then asked for suggestions of other features that would be helpful in creating a user-friendly data set. Their responses are detailed below in no particular priority ranking.

Transcripts

There was a unanimous desire for audio transcripts. One participant commented that transcribing the interactions around speech "opens up extreme richness across the board for everything we're talking about." One could even imagine creating dynamic transcripts that permit users to toggle back and forth between segments of the spoken word and the corresponding video clips.

Simultaneous Viewing

Participants agreed that being able to display the time-synchronized clips from all the rooms in a home on a tiled or split screen simultaneously, similar to CCTV screens in a security office, would be very valuable, especially because it would permit more holistic coding.

Transitions

It was proposed that "transitions" from one activity to another are important. While some transitions may occur within the same video clip, it would be necessary to have a way to easily access and link adjacent clips to observe transitions that span multiple clips.

Linking Events Across Time

The more general ability to link videos that are connected "forwards and backwards in time" by the participants themselves was also suggested. For example, one might want to link a clip in which someone mentions that they need to perform a certain activity with the clip where the activity was actually performed.

Surroundings and Objects

Having a sense of the material objects in the home such as artwork, books, newspapers, or artifacts could be important as they relate, for example, to racial pride, culturally grounded practices, or resources for children. Items that are visible from the video clips could be supplemented by data from the hand-held camcorder used during the home visit, an earlier step in the process of becoming acquainted with potential NJFS families.

Measuring Proximity

Measuring the proximity or distance between individuals in a clip would be desirable to determine the physical "closeness" of children with adults and adults with each other.

Noise levels

There was a brief mention about the possibility of measuring noise levels in a clip, which could be accomplished using the decibel readings from the metadata.

Dynamic Data Archive

As a condition for using the NJFS data, researchers could be required to archive their more granular codes on the project's server. The research community could then have access to them, see what has already been done, and build on them over time. Doing so would allow the NJFS data resource to take on a "self-propelling, life of its own" and be "truly transformative." It could even be possible for individuals to suggest additional macrocodes that early childhood researchers would find useful.

Editable Database

A related suggestion pertained to creating an "editable database." As people work with the data, they might add codes to clips that other researchers can pay attention to or not. It was argued that older data gets better when more eyes are on it and as different analytic perspectives are brought to bear. With more descriptions and analyses, one has an ever-richer characterization of what's happening. If everybody who encounters the data has access to everybody else's suggestions, one keeps building a cumulative science. This would not be a trivial undertaking, because it involves issues of quality assurance and data storage, among others. But the result "would benefit generations of people."

Searchable Database

Given the large number of video clips, most participants felt it would be necessary to build a searchable database that would permit a researcher to find easily and quickly on those clips of greatest interest to the research question at hand. To make this concept more concrete, the home buyers' website www.realtor.com was offered as an example. Homebuyers can use filters such as "Price," "Number of bedrooms," and "Zip code" to narrow their search options.

Artificial Intelligence

There are creative opportunities for using artificial intelligence (AI) for coding NJFS video clips. The groups felt there was much potential here: "AI is powerful." Indeed, it's a necessity given the staggering amount of data; there is just too much to code by hand. There are automated programs for coding emotional expression. What kinds of tools are available to capture motion, speech, activity? Distinguishing between "something happening" in a room and "nothing happening" can be done automatically with a decent level of precision. One participant noted that the geometry of the visuals in many studies is a computer vision challenge. With the NJFS, however, the stable placement of cameras over time is a real plus. The Princeton NJFS team is in conversation with faculty in Computer Science who specialize in computer vision and speech recognition.

APPENDIX A: Focus Group Participants

New Jersey Families Study

Focus Group #1
Friday, December 4, 3:00-4:30pm EST

Vivian L. Gadsden
University of Pennsylvania

Roberta M. Golinkoff
University of Delaware

Stephanie M. Jones
Harvard University

Casey Lew-Williams
Princeton University

Catherine Tamis-Lemonda
New York University

Focus Group #2
Monday, December 7, 3:00-4:30pm EST

Jeanne Brooks-Gunn
Columbia University

Caitlin M. Fausey
University of Oregon

Ellen Galinsky
Bezos Family Foundation

Iheoma U. Iruka
University of North Carolina

Geoffrey Raymond
University of California

NJFS Princeton Group

Dawn Crossland-Sumners
Focus group moderator

Thomas Espenshade
Project director

Jeffrey Himpele
Co-Principal Investigator

Cecilia Kim
Rapporteur

Boriana Pratt
Data manager/statistical programmer

Daniel Veith
IT Support

APPENDIX B: Focus Group Participant Biographies

Jeanne Brooks-Gunn

Prof. Jeanne Brooks-Gunn is a developmental psychologist and the Virginia and Leonard Marx Professor of Child Development at Teachers College, Columbia University. Her research interests focus on language development and her publications include "Early child development in the 21st Century: Profiles of current research initiatives" and "First-year maternal employment and child development in the first seven years."

Caitlin M. Fausey

Prof. Caitlin Fausey is an assistant professor in the department of Psychology at the University of Oregon. She is currently working with "The Learning Lab," a study analyzing early childhood development and learning using audio recording and video footage. Her research interests focus on early sensory input, and her publications include "From faces to hands: Changing visual input in the first two years" and "Contributions of head-mounted cameras to studying the visual environments of infants and young children."

Vivian L. Gadsden

Prof. Vivian Gadsden is the William T. Carter Professor of Child Development and Education, Director of the National Center on Fathers and Families, and Associate Director of the National Center on Adult Literacy at the Graduate School of Education at the University of Pennsylvania. Her research interests focus on cultural and social factors affecting learning and literacy across the life-course and within families. Her publications include "'Urban' Schooling and 'Urban' Families: The Role of Context and Place" and "Parenting Matters: Supporting Parents of Children Ages 0-8."

Ellen Galinsky

Ellen Galinsky is an expert on early childhood learning and intervention. She is the Chief Science Officer at the Bezos Family Foundation where she also serves as the director of "Mind in the Making," an early childhood learning program. She is also the president of the Families and Work Institute. Her research interests focus on child care, parent-professional relationships, parental development, work-family issues, and youth voice. Her books include *Mind in the Making* and *Ask the Children*.

Roberta M. Golinkoff

Prof. Roberta Golinkoff holds the Unidel H. Rodney Sharp Chair in the School of Education at the University of Delaware. She is also a professor in the Departments of Psychological and Brain Sciences and Linguistics and Cognitive Science. Her research interests focus on language and development of early spatial knowledge. Her publications include "How Babies Talk" and "Becoming Brilliant: What Science Tells Us About Raising Successful Children."

Iheoma U. Iruka

Prof. Iheoma Iruka is a Research Professor of Public Policy and the Founding Director of the Equity Research Action Coalition at the Frank Porter Graham Child Development Institute, University of North Carolina at Chapel Hill. Her research interests focus on racial equity research, programming, and policy, especially for Black children, and her publications include "Profiles of parenting for low-income families and links to children's preschool outcomes" and "Early Steps to School Success (ESSS): Examining pathways linking home visiting and language outcomes."

Stephanie M. Jones

Prof. Stephanie Jones is Professor of Education at the Harvard Graduate School of Education. Her research interests focus on child emotional and social development that underlie and support learning. Her publications include "The Leading Edge of Early Childhood Education" and "The infant–toddler social and emotional assessment (ITSEA): Factor structure, reliability, and validity."

Casey Lew-Williams

Prof. Casey-Lew Williams is a Professor of Psychology at Princeton University, focusing on early childhood learning and development. He is the co-director of the "Princeton Baby Lab," which studies patterns in learning behavior of babies. His research interests focus on how parent-child interaction helps child cognition in the short and long term and the effects of interaction on the learning process. His publications include "Bilingual infants control their languages as they listen" and "Mothers consistently alter their unique vocal fingerprints when communicating with infants."

Geoffrey Raymond

Prof. Geoffrey Raymond is a Professor of Sociology at the University of California, Santa Barbara. His research interests include conversation analysis and organization of interaction in socializing children, the role of talk-in-interaction in the organization of institutions, and qualitative research methods. His publications include "Talk and interaction in social research methods" and "The question of units for language, action and interaction."

Catherine Tamis-Lemonda

Prof. Catherine Tamis-Lemonda is a Professor of Applied Psychology at New York University. She is working on a study called the "Play and Language Lab," that explores early childhood learning and how various cultural and social conditions affect learning. Her research interests focus on how babies navigate and learn language from their environment, as well as developmental cascades. Her publications include "Fathers and mothers at play with their 2-and 3-year-olds: Contributions to language and cognitive development" and "Maternal responsiveness and cognitive development in children."

APPENDIX C: Proposed Metadata Codes

Household-specific information

1. Household Identification Number
2. Two-week video recording start date (Day/Month/Year)
3. Two-week video recording start day (Day of week: Mon, Tues, etc.)

Camera-specific information

4. Camera number
5. Camera location (living room, kitchen, etc.)

Clip-specific information

6. Measures of decibel readings (e.g., min, max, measures of central tendency—median, mean) and the time when the maximum was recorded.
7. Duration of clip in minutes or seconds and in discrete categories.
8. Day of week (Mon, Tues, ..., Sat, Sun)
9. Weekday or Weekend
10. Week 1 or Week 2
11. Time of day: Morning (6am – noon); Afternoon (noon – 6pm); Evening (6pm – midnight); Night (midnight to 6am)

Note: Items 8-11 refer to the time an individual video clip begins.