

Running head: PEDAGOGICAL QUESTIONS IN CHILDES

Pedagogical Questions in Parent-Child Conversations

Yue Yu¹² (yue.yu@rutgers.edu), Elizabeth Bonawitz² (elizabeth.bonawitz@rutgers.edu),

Patrick Shafto¹ (patrick.shafto@gmail.com)

Department of ¹Math and Computer Science & ²Psychology,

Rutgers University - Newark

Author Note

This study is supported by NSF CAREER, DRL-1149116 to P.S and NSF Award SMA-1640816 to E.B. and P.S. We thank Reham Bader, Jack Fredricks, Victoria Golinski and Rosa Lasso for coding. We thank four anonymous reviewers for providing constructive comments and suggestions. Correspondence concerning this article should be addressed to Yue Yu, Department of Mathematics & Computer Science, Rutgers University-Newark, NJ 14850. Email: yue.yu@rutgers.edu.

Word Count: 5888

Abstract

Questioning is a core component of formal pedagogy. Parents commonly question children, but do they use questions to teach? Here we define “pedagogical questions” as questions for which the questioner already knows the answer and intended to help the questionee learn. We investigate the frequency and distribution of pedagogical questions from parent-child conversations documented in the CHILDES database. By analyzing 2166 questions from 166 mother-child dyads and 64 father-child dyads, we find that pedagogical questions are commonplace during day-to-day parent-child conversations, and vary based on child’s age, family environment, and historical era. The results serve as a first step towards understanding the role of parent-child questions in facilitating children’s learning.

Keywords: informal pedagogy, pedagogical questions, parent-child conversation, socioeconomic status, CHILDES.

Pedagogical Questions in Parent-Child Conversations

To question well is to teach well.

- Henry Barnard, 1860, American Journal of Education

Asking and answering questions has been seen as a core component of teaching and learning at least since the days of Socrates. Teachers ask ~400 questions per day in a typical classroom setting (Gall, 1970), serving various functions from checking class work to motivating thinking (Black, 2001). Research in education suggests that questioning leads to improvements in academic outcomes (Redfield & Rousseau, 1981; Von Secker, 2002; Wise & Okey, 1983). Research in cognitive development has shown that parents engage in question-asking throughout childhood (Ervin-Tripp & Miller, 1977), starting by at least as early as five-months-old (Bornstein et al., 1992). Questions seem to support language learning (Blewitt, Rump, Shealy, & Cook, 2009; Sénéchal, 1997) as well as general content learning (Haden, Cohen, Uttal, & Marcus, 2015). Thus, past research has demonstrated that questions are both ubiquitous in early childhood and also a potential key to children's learning and development.

However, for a learner, not all questions are created equal. Research has estimated that around 80% of questions asked in classrooms are for teaching (Gall, 1970), among them a majority are prompts for facts and a minority are prompts to elicit thinking (Siraj-Blatchford & Manni, 2008). These questions by a (presumably) knowledgeable teacher to the less knowledgeable learners are unusual relative to the kinds of questions we most commonly think about in day-to-day conversation. Most commonly, questions are a tool for eliciting information from others ("information-seeking questions") or to make a point ("rhetorical questions") (Searle, 1969; Shatz, 1978). However, "pedagogical questions"—questions that are asked by a knowledgeable individual whose goal is teaching—are unique from other kinds of questions.

Whereas information-seeking questions are asked to elicit unknown information from the questionee, and rhetorical questions are questions to which both or neither of the parties know the answer, pedagogical questions are questions asked by the person who knows the answer (or might reasonably be expected to know the answer) to someone who may not know the answer, with the goal of eliciting learning.

Do parents employ pedagogical questions? We know that parent-child conversation plays a vital role in facilitating children's thinking and learning (Crowley et al., 2001; Frazier, Gelman, & Wellman, 2009), and questioning is prevalent in parental discourse from early on in development (Bornstein et al., 1992; Kurkul & Corriveau, in press). Therefore, it becomes important to ask whether pedagogical questions are commonplace. Previous functional and syntactic categorizations of questions have identified types of questions that may relate to facilitating learning, such as *wh*- questions, open-ended questions, test questions, and knowledge questions (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956; Chouinard, Harris, & Maratsos, 2007; Ervin-Tripp, 1970; Olsen-Fulero & Conforti, 1983; Rowe, Leech, & Cabrera, 2016; Shatz, 1979). However, none of these categories accurately capture the essence of questions used for teaching. For example, *wh*- and open-ended can be used to seek information from children (e.g., "What do you want to play?"), whereas test and knowledge questions can be used to check children's memory without trying to teach new information (e.g., "Do you remember that?"). Therefore, in this paper we take a new perspective and analyze parent-child questions by focusing on the knowledge state and the intention of the parent.

Pedagogical questions support learning

Theories on informal pedagogy (Csibra & Gergely, 2009; Tomasello, 1999) may help us understand how parents help children learn through informal instruction. Infants and young

children are sensitive to cues that suggest an adult's intention to teach, such as joint attention, child-directed speech, name-calling, etc. (Csibra & Gergely, 2009). Older children also consider whether the teacher is knowledgeable when drawing inferences in teaching contexts (Bonawitz et al., 2011; Gonzalez, Shafto, Bonawitz, & Gopnik, 2012; Harris & Corriveau, 2011; Koenig, Clément, & Harris, 2004; Shafto, Goodman, & Frank, 2012). These “pedagogical cues” engage strong inferences, which facilitates learning in various domains (Bonawitz et al., 2011; Buchsbaum, Gopnik, Griffiths, & Shafto, 2011; Butler & Markman, 2014; Nielsen, 2006; Sage & Baldwin, 2010; Topál, Gergely, Miklósi, Erdőhegyi, & Csibra, 2008; Vredenburg, Kushnir, & Casasola, 2014).

Like pedagogical instruction, pedagogical questions imply an opportunity for learning for the questionee. An ignorant person asking “What does that button do?” is likely simply seeking information about the button. However, when the exact same question is asked by a person who is assumed to already know the answer, more can be inferred by the questionee: The questioner is probably trying to convey something worth learning about the button. Our recent experiments have shown that when presented with a novel toy, 4- and 5-year-old children indeed learned and explored more after being asked “What does the button do?” from a knowledgeable teacher, rather than after being asked the exact same question from an ignorant bystander (Yu, Landrum, Bonawitz, & Shafto, under review). Children also take the knowledge state of a questioner into account when deciding whether to revise a previously stated belief (Bonawitz et al., under review). Additionally, children assume that a teacher, but not a bystander, will ask them about something the questioner already knows: When an adult asked “Where is X?” during an ongoing pedagogical interaction with a 2-year-old, the toddler was more likely to assume that “X” referred to an object known to the adult and the adult was trying to draw her attention to that

object; whereas when an adult asked “Where is X?” while doing things alone, the toddler was more likely to assume that “X” referred to an object unknown to the adult and the adult is searching for that object (Grosse & Tomasello, 2012). These findings suggest that young children associate the knowledge state of a questioner with his or her pedagogical intent, which has implications for their learning.

Pedagogical questions may also differ from pedagogical instructions in important ways: While pedagogical instructions facilitate learning of what is being instructed, they also elicit the learner’s inference that the teacher is purposefully choosing what to instruct and not to instruct, so what is not chosen is probably unimportant and need not be explored (Shafto, Goodman, & Griffiths, 2014). These inferences have been shown to lead to less exploration and further learning when the environment afford more learning opportunities than what was instructed (Bonawitz et al., 2011; Shneidman, Gweon, Schulz, & Woodward, 2016). However, the constraints on exploration and further learning is not observed when pedagogical instructions are reframed as pedagogical questions (Yu et al., under review), suggesting that the choice of pedagogical method has implications on children’s learning.

Contextual factors

Several factors might influence the use of pedagogical questions. For example, past research has found that the type of questions parents ask children changes with children’s age (Kuchirko, Tamis-LeMonda, Luo, & Liang, 2015; Snow et al., 1976). Previous studies have also identified differences between mother-child and father-child conversations (Gleason, 1975; Rowe, Coker, & Pan, 2004; Walker & Armstrong, 1995), as well as an effect of father’s presence on mother-child interaction (Clarke-Stewart, 1978). Therefore, in our analysis we included children’s age and the presence of mothers and fathers as possible predictors of questioning behavior.

Pedagogy also varies across culture and time (Shneidman, Gaskins, & Woodward, 2015), and the types of questions parents ask children differ based on family environment (Hoff, 2003; Snow et al., 1976). For example, during free play, mothers from academic middle-class families posed more *wh*-questions and less yes-no questions than mothers from working-class and lower middle-class families (Snow et al., 1976), and the use of *wh*-questions in father-child interactions were shown to predict children's language development in low-income families (Rowe et al., 2016). Based on these findings, we have included family socioeconomic status (SES) as well as location and time of data collection as possible predictors of questioning behavior.

CHILDES Analyses

To address whether and how parents use pedagogical questions, we analyzed parent-child conversations from the CHILDES database (MacWhinney & Snow, 1990). We coded parents' spontaneous child-directed questions into three categories: Pedagogical questions are those for which parents know the answer and want children to learn; information-seeking questions are those for which parents do not know the answer and seek the answer from the child; and rhetorical questions are those not intend to be answered, either because there is no obvious answer or because it is assumed that both parties already know the answer. We also recorded parents' gender, child's age and gender, context of the conversation, measurements of family environment, and time and location of data collection, so as to examine how parents' questioning behavior may vary based on these factors.

Method

Sample

We searched the entire CHILDES database (MacWhinney & Snow, 1990) for transcripts that meet the following nine criteria:

1. The transcript was in English.
2. The conversation took place at home.
3. At least one parent and one child participated in the conversation, and no one outside the immediate family (interviewer, grandparents, relatives, friends, etc.) was involved.
4. The target child was between 2 and 6 years of age.
5. The conversation represented everyday talk, and was not a purposeful conversation such as an interview.
6. The transcripts for parents' and child's speech were not separated.
7. The transcript used punctuation marks.
8. The transcript contained at least three questions between a parent and a child.
9. If there were multiple transcripts for a same parent-child dyad (such as in longitudinal studies), we only used the first (earliest) transcript that meets all other criteria.

According to these criteria, 185 transcripts were collected from 27 studies, which included 166 mother-child conversation samples and 64 father-child conversation samples (Figure 1). Detailed information for each study is listed in Table 1. For each study, we recorded the time and location of data collection. Because detailed information on data collection was not available for all studies, we specified time at the level of decade (1970s, 1980s, 1990s, 2000s), and specified location at the level of country (USA, UK). We recorded family SES for studies that reported a homogenous sample (e.g., all participants were from working-class or middle-class families), and for studies that specified SES for each individual family. We did not record SES if the study only reported SES of the neighborhood from which the families were recruited, but did not report SES of the actual families that participated. Whenever possible, we also recorded sampling and procedural information to control for possible differences across studies. These include whether

the study used random sampling or convenience sampling (e.g., researcher's own child), number of children we included from the study, the settings of the recordings, the instructions given to the parents, and the length of recording (Table 1). Two of these factors—sampling method of the study and the number of children we included from the study—were easy to quantify and were available for all studies, so we added them as study-level control variables in our statistical analyses. Neither predicted parents' questioning behavior or changed the significance of other predictors (Table 3).

Coding Procedures

Step-by-step coding instructions and the training process of the coders are detailed in the Supplementary Material. For each transcript, one coder first recorded the target child's gender, age, and the conversation partners from the heading of the transcript. The coder then derived the total number of the parent's statements and questions in the transcript using CLAN, the programming tool provided by the database (MacWhinney, 2000). The frequency of questioning (per 100 statements) was calculated by dividing the number of questions by the total number of statements, and then multiplying by 100.

Following the recording of this initial information, two coders who were blind of the hypotheses independently coded the first 10 questions that 1) ended with a question mark; 2) were asked by the parent; 3) were directed towards the target child; and 4) did not contain missing words. Out of the 230 conversation samples, 30 contained less than 10 such questions (with a minimum of 3 questions), and for those transcripts all questions were coded. Each question was coded according to the coding scheme listed in Table 2: The coders first determined whether the question was a pedagogical question, information-seeking question, or rhetorical question based on the knowledge state and the intention of questioner. They then

assigned the questions into one of eight subcategories: The pedagogical questions were further distinguished based on whether they were intended to teach generic or specific knowledge (Gelman, Goetz, Sarnecka, & Flukes, 2008), and the information-seeking and rhetorical questions were further distinguished based on their functions (Olsen-Fulero & Conforti, 1983). In cases when parent's knowledge state and intention were difficult to judge by the question itself, coders would then refer to 1) contexts before and after the question, 2) the linked audio or video clips if they were available. Since transcripts do not fully capture the history and nonverbal aspects in parent-child interactions, and audio or video clips were not always available, some questions could remain ambiguous despite coders' best effort. In Supplementary Material we provided an estimation of the frequency of these ambiguous situations, and explained why they should have limited impact on our final results. Overall the inter-rater reliability was high, which shows the majority of questions can be reliably categorized based on our coding scheme: When only the three major categories were considered, Cohen's $\kappa = .830$; When all eight subcategories were considered, Cohen's $\kappa = .832$. Inconsistent codes were reviewed and resolved by a third coder.

To ensure that questions for which parents knew the answer of were indeed pedagogical, a fourth independent coder (who was not involved in the initial coding, but familiar with research on informal pedagogy) checked a subsample of questions coded as pedagogical ($n = 132$) to see "if the questions can be interpreted as intended to teach, where teaching is broadly defined as any endeavor to help children learn". Among the 132 questions, 123 (93.2%) were coded as clearly intended to teach, and the rest 9 questions (6.8%) were ambiguous instances in which the parent may be intended to help the child learn, but the transcript did not provide enough context to accurately judge the parent's intentions. We take these findings as evidence that our coding of

pedagogical questions indeed captured parents' pedagogical intent. Details of the procedure and results of this additional coding are described in the Supplementary Material.

For each question, the two initial coders also decided whether the child responded to the question (repeated the question or provided an answer), whether the parent followed-up the question (repeated the question, provided an answer, or added relevant information), or whether the next line is irrelevant (a different person spoke or the topic was changed). Inter-rater reliability was high (Cohen's $\kappa = .846$), and inconsistent codes were reviewed and resolved by the third coder.

After coding all 10 questions, the first two coders also determined whether the main context of the conversation was 1) over meal (lunch or dinner), 2) during free play (e.g., playing with toys, book reading, spontaneous play or talk), or 3) during daily routines (e.g., taking a bath, preparing to leave).

Data analysis

All data was entered and analyzed in IBM SPSS 22. Given the multilevel structure of our sample (parent-child dyads were nested under studies), we used multilevel mixed-effects linear regression models as the analytical tool for omnibus tests (advantages of using multilevel models to analyze archival data are detailed in the Supplementary Material). Fisher's exact test was used for comparisons of frequencies. An α level of .05 (two-tailed) was used for all tests.

Results

Table 2 shows the average proportion of each category and subcategory of questions in mother-child and father-child conversations. Mothers and fathers did not differ in the proportion of pedagogical and information-seeking questions they asked, although there were differences in rhetorical questions and subcategories of information-seeking questions (Table 2). To examine

the contributing factors to parents' questioning behavior, we started with a multilevel model that includes all conversation-level, child-level, and study-level predictors as independent variables, to predict the proportion of pedagogical questions (Table 3, Model 1). Results showed that the child's age, the SES of the family, and the time of data collection had significant effects on parents' pedagogical questions towards children. We then reduced Model 1 by excluding all non-significant predictors. The revised model (Model 2) was superior to Model 1 by the criterion of restricted log likelihood and Bayesian Information Criterion (Table 3), and was robust when cross-validation was applied (details see Supplementary Material).

Model 2 showed that the proportion of pedagogical questions decreased with children's age (Figure 2), $B = -.004$, $p = .038$. Parents asked a larger proportion of pedagogical questions to toddlers (24-36 months; $n = 72$; $M = 34\%$) than to preschoolers (36-72 months; $n = 94$; $M = 25\%$), $t(179.4) = 2.43$, $p = .016$, $d = 0.33$. When we separated subcategories of generic and specific pedagogical questions, the age trend was significant for the generic questions, $B = -.004$, $p < .001$, but not the specific questions, $B = .000$, $p = .86$. Additionally, the proportion of pedagogical questions was higher in families noted as middle-class ($n = 30$; $M = 55\%$) than in families noted as working-class ($n = 85$; $M = 19\%$), $B = .261$, $p = .003$, and this was true for the specific questions, $B = .245$, $p = .007$, but not general questions, $B = .028$, $p = .57$. Finally, the proportion of pedagogical questions increased with historical era, $B = .063$, $p = .038$, with parents from the 2000s asking a larger proportion of pedagogical questions ($n = 14$; $M = 42\%$) than parents from the 1970s ($n = 64$; $M = 26\%$), $t(76) = 2.30$, $p = .024$, $d = 0.68$. Again, this was true for the specific questions, $B = .056$, $p = .030$, but not general questions, $B = .008$, $p = .63$. Further analysis showed no significant two-way interaction effects between child's age, SES, and historical era.

To examine whether child's age, SES, and historical era also have an impact on the other types of questions, we applied Model 2 to the proportion of information-seeking and rhetorical questions, as well as parents' overall frequency of questions (per 100 statements). Results showed that the proportion of information-seeking questions increased with children's age, $B = .005$, $p = .007$, and decreased with historical era, $B = -.070$, $p = .032$. The proportions of information-seeking and rhetorical questions were also lower in middle-class families than in working-class families (Figure 3), $Bs < -0.110$, $ps < .066$. The overall frequency of questions (per 100 statements) was not predicted by child's age or historical era, $ps > .32$, but it was higher for middle-class parents compared to working-class parents, $B = .137$, $p = .011$. To examine if the difference in overall frequency of questions is responsible for the different proportions of question types observed between middle-class and working-class families, we estimated the frequencies of different question types by multiplying the proportion of the question types by the overall frequency of questions. According to this estimate, middle-class parents asked pedagogical questions more frequently than working-class parents, $B = .107$, $p = .011$, and the frequency of information-seeking and rhetorical questions did not differ between middle-class and working-class parents, $Bs < .017$, $ps > .29$. Therefore, the higher proportion of pedagogical questions in middle-class families was driven by higher frequencies of these questions, and not by lower frequencies of information-seeking or rhetorical questions.

We further examined whether the same predictors are associated with mothers' and fathers' questioning. The proportion of mothers' pedagogical questions decreased with age, $B = -.005$, $p = .009$, but age was not a significant predictor for fathers' pedagogical questions, $B = .001$, $p = .85$ (Figure 2). SES is strongly associated with both mothers' and fathers' questions: Middle-class mothers and fathers asked a higher proportion of pedagogical questions than working-class

mothers and fathers, $ts > 43.02$, $ps < .001$ (Figure 3). Mothers' pedagogical questions increased with historical era, $B = .054$, $p = .049$, with mothers from the 2000s asking a higher proportion of pedagogical questions than mothers from the 1970s and 1980s, $ts > 2$, $ps < .05$, $ds > 0.7$. Fathers' pedagogical questions also show an increasing trend with historical era (Figure 4), although the trend was non-significant, $B = .045$, $p = .12$. Finally, an interaction effect between parents' gender and the presence of other parent was observed for the proportion of pedagogical questions, $F(1, 226) = 4.36$, $p = .038$ (Figure 5). Mothers asked more pedagogical questions when a father is noted as present compared to not present, $B = .083$, $p = .045$, but fathers' pedagogical questions did not differ based on mothers' presence, $B = -.172$, $p = .17$.

To better understand the nature of parents' pedagogical questions, we also looked at the responses and follow-ups of these questions. Here the follow-up is defined as a relevant statement immediately after the initial question, and can be a repeat of the question, an answer to the question, or the addition of question-relevant information. Overall children responded to a similar proportion of pedagogical questions (47%) and information-seeking questions (46%), Fisher's exact $p = .70$. However, parents were more likely to follow-up a pedagogical question (23%) than an information-seeking question (17%, $p = .004$) or a rhetorical question (16%, $p = .043$). Parents were also more likely to follow-up pedagogical questions for toddlers (26%) than for preschoolers (19%), $p = .036$, despite that both toddlers (45%) and preschoolers (49%) responded to a same amount of pedagogical questions, $p = .34$. These results suggest that pedagogical questions are different in kind from information-seeking or rhetorical questions: They involve greater degrees of interaction between parent and child, which appear to reflect differences in parents' behavior rather than children's responsiveness.

Discussion

Educators use questions for which they already know the answer to guide students' learning. Our results suggest that parents do as well. Moreover, the proportion of pedagogical questions parents ask their children varies based on several factors, including children's age, family SES, and the historical era. Mothers' tendencies to ask pedagogical questions also differed based on whether a father is present during the conversation.

The finding that parents' pedagogical questioning differed by children's age is consistent with previous research showing that parents adjust utterance in general (Snow, 1972), and questions in particular (Kuchirko et al., 2015), with regard to the age of their children. Specifically, we have shown that the proportion of pedagogical questions, especially those asking about general kinds, was higher in parents' questions towards their toddlers than their preschoolers. These developmental changes, together with evidence that American mothers ask pedagogical questions even to 5-month-old infants (Bornstein et al., 1992), may shed light on the nature of pedagogical questions. First, parents' use of questions to teach, and the benefits it brings to children's learning, may occur well before children can understand or answer these questions. Indeed, research has shown that even infants could be sensitive to others' pedagogical intent and knowledge states (Csibra & Gergely, 2009; O'Neill, 1996), which raises the possibility that pedagogical questions could help drive infants' attention and serve as language input without the expectation of being answered verbally. Acoustic features like interrogative prosody have been suggested to serve as a cue to facilitate learning (Bornstein & Lamb, 2002), raising a potentially relevant avenue for understanding the mechanism behind how early pedagogical questioning are used to teach. Second, the decreasing trend of parents' pedagogical questioning from toddlerhood to preschool years is in contrary to children's increasing abilities and needs to learn. It is possible that parents are changing tools to stir learning—older children may require

fewer pedagogical cues to indicate an opportunity to learn, so parents may switch to more direct forms of teaching. It is also possible that because of an increasing ability to respond to information-seeking questions, older children are asked more information-seeking questions instead of pedagogical questions.

Consistent with previous studies (e.g., Clarke-Stewart, 1978; Hoff, 2003; Snow et al., 1976), we show that family environment plays a major role in how parents ask children questions. Mothers and fathers from working-class families ask less than half as many pedagogical questions as mothers and fathers from middle-class families. Additionally, mothers also ask more pedagogical questions when a father is present. Further research is needed to identify the family dynamics underlying question asking in different social groups, and their implications on children's learning and school readiness. Nonetheless, the current results add to a growing body of research suggesting the importance of examining the *quality* of parent-child interactions, over and above *quantity*, on children's cognitive development (Hirsh-Pasek et al., 2015).

Finally, we found parents' questioning practices to change over recent history, with millennial mothers asking significantly more pedagogical questions than mothers from the 1970s and 80s. This effect may be related to the historical trend of parents becoming older and more educated (Pew Research Center, 2010). It points to the importance of situating research in the historical contexts when studying parenting practices in general, and questioning behavior in particular.

The sample for our study comes from the CHILDES database, so the scope is limited to information made available, and for the available variables the data were not balanced. Random-assignment experiments are needed to confirm our findings, and to extend them to cover other important factors. For example, although we did not find differences in questioning between

parents from USA and UK, these results may not hold universally. In fact, mothers from western societies may be unique in their strong tendencies to ask questions to infants (Bornstein & Lamb, 2002), so future study need to verify the results in other cultures.

In conclusion, this study builds upon and extends an accumulating literature on the role of informal pedagogy in children's learning. Parents may teach not only through direct instruction, but also through asking questions for which they know the answer. Our results suggest that parents do consistently use pedagogical questions in their day-to-day conversations, and the relative frequencies of these questions depend on children's age, family environment, and historical context. A critical next step, then, is to evaluate children's inferences from, and learning outcomes following, pedagogical questions. Doing so will help bridge theories of pedagogy and inquiry-based learning, and bring us closer to understanding how to question well.

References

- Black, S. (2001). Ask me a question: How teachers use inquiry in a classroom. *American School Board Journal*, 188(5), 43-45.
- Blewitt, P., Rump, K. M., Shealy, S. E., & Cook, S. A. (2009). Shared book reading: When and how questions affect young children's word learning. *Journal of Educational Psychology*, 101(2), 294-304. doi:10.1037/a0013844
- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). Taxonomy of educational objectives, handbook I: The cognitive domain. New York: David McKay Co Inc.
- Bohannon III, J. N., & Marquis, A. L. (1977). Children's control of adult speech. *Child Development*, 48(3), 1002-1008. doi:10.2307/1128352
- Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition*, 120(3), 322-330. doi:10.1016/j.cognition.2010.10.001
- Bonawitz, E., Shafto, P., Yu, Y., Gonzalez, A., Bridgers, S., & Gopnik, A. (under review). Children change their answers in response to neutral follow-up questions by a knowledgeable asker.
- Bornstein, M. H., & Lamb, M. E. (2002). *Development in infancy: An introduction*. New York: McGraw-Hill.
- Bornstein, M. H., Tal, J., Rahn, C., Galperin, C. Z., Pêcheux, M.-G., Lamour, M., . . . Tamis-LeMonda, C. S. (1992). Functional analysis of the contents of maternal speech to infants of 5 and 13 months in four cultures: Argentina, France, Japan, and the United States. *Developmental Psychology*, 28(4), 593. doi:10.1037/0012-1649.28.4.593

- Braunwald, S. R. (1976). Mother-child communication: The function of maternal-language input. In R. Von Raffler-Engel (Ed.), *Child language-1975* (pp. 28-50). Great Britain: International Linguistics Association.
- Buchsbaum, D., Gopnik, A., Griffiths, T. L., & Shafto, P. (2011). Children's imitation of causal action sequences is influenced by statistical and pedagogical evidence. *Cognition, 120*(3), 331-340. doi:10.1016/j.cognition.2010.12.001
- Butler, L. P., & Markman, E. M. (2014). Preschoolers use pedagogical cues to guide radical reorganization of category knowledge. *Cognition, 130*(1), 116-127. doi:10.1016/j.cognition.2013.10.002
- Chouinard, M. M., Harris, P. L., & Maratsos, M. P. (2007). Children's questions: A mechanism for cognitive development. *Monographs of the Society for Research in Child Development, i*-129.
- Clarke-Stewart, K. A. (1978). And daddy makes three: The father's impact on mother and young child. *Child Development, 466*-478.
- Crowley, K., Callanan, M. A., Jipson, J. L., Galco, J., Topping, K., & Shrager, J. (2001). Shared scientific thinking in everyday parent-child activity. *Science Education, 85*(6), 712-732.
- Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*(4), 148-153. doi:10.1016/j.tics.2009.01.005
- Demetras, M. (1986). *Working parents' conversational responses to their two-year-old sons*. University of Arizona.
- Demetras, M. (1989). *Changes in parents' conversational responses: A function of grammatical development*. Paper presented at the ASHA, St. Louis, MO.
- Demetras, M., Post, K. N., & Snow, C. E. (1986). Feedback to first language learners: The role

- of repetitions and clarification questions. *Journal of Child Language*, 13(02), 275-292.
doi:10.1017/S0305000900008059
- Demuth, K., Culbertson, J., & Alter, J. (2006). Word-minimality, epenthesis and coda licensing in the early acquisition of English. *Language and Speech*, 49(2), 137-173.
doi:10.1177/00238309060490020201
- Dickinson, D. K., & Tabors, P. O. (2001). *Beginning literacy with language: Young children learning at home and school*. Baltimore: Paul Brookes Publishing.
- Ervin-Tripp, S. (1970). Discourse agreement: How children answer questions. *Cognition and the development of language*, 79-107.
- Ervin-Tripp, S., & Miller, W. (1977). Early discourse: Some questions about questions. In M. M. Lewis & L. A. Rosenblum (Eds.), *Interaction, conversation, and the development of language*. New York: Wiley.
- Feldman, A., & Menn, L. (2003). Up close and personal: A case study of the development of three English fillers. *Journal of Child Language*, 30(04), 735-768.
doi:10.1017/S0305000903005774
- Frazier, B. N., Gelman, S. A., & Wellman, H. M. (2009). Preschoolers' search for explanatory information within adult-child conversation. *Child Development*, 80(6), 1592-1611.
- Gall, M. D. (1970). The use of questions in teaching. *Review of Educational Research*, 707-721.
- Gelman, S. A., Goetz, P. J., Sarnecka, B. W., & Flukes, J. (2008). Generic language in parent-child conversations. *Language Learning and Development*, 4(1), 1-31.
- Gleason, J. B. (1975). Fathers and other strangers: Men's speech to young children. In D. P. Dato (Ed.), *Developmental psycholinguistics: Theory and applications* (pp. 289-297). Washington, DC: Georgetown University Press.

- Gleason, J. B. (1980). The acquisition of social speech: Routines and politeness formulas. In H. Giles, W. Robinson, & P. Smith (Eds.), *Language: Social psychological perspectives* (pp. 159-167). Oxford, UK: Pergamon.
- Gonzalez, A., Shafto, P., Bonawitz, E., & Gopnik, A. (2012). Is that your final answer? The effects of neutral queries on children's choices. *Proceedings of the thirty-fourth cognitive science society*.
- Grosse, G., & Tomasello, M. (2012). Two-year-old children differentiate test questions from genuine questions. *Journal of Child Language*, 39(01), 192-204.
- Haden, C. A., Cohen, T., Uttal, D. H., & Marcus, M. (2015). Building learning: Narrating and transferring experiences in a children's museum. In D. Sobel & J. Jipson (Eds.), *Cognitive Development in Museum Settings: Relating Research and Practice* (pp. 84-103). New York, NY: Routledge.
- Harris, P. L., & Corriveau, K. H. (2011). Young children's selective trust in informants. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1567), 1179-1187. doi:10.1098/rstb.2010.0321
- Hayes, D. P., & Ahrens, M. G. (1988). Vocabulary simplification for children: A special case of 'motherese'? *Journal of Child Language*, 15(2), 395-410. doi:10.1017/S0305000900012411
- Henry, A. (1995). *Belfast English and Standard English: Dialect variation and parameter setting*. New York: Oxford University Press.
- Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., . . . Suma, K. (2015). The contribution of early communication quality to low-income children's language success. *Psychological Science*, 0956797615581493.

- Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development, 74*(5), 1368-1378.
- Howe, C. (1981). *Acquiring language in a conversational context*. New York: Academic Press.
- Koenig, M. A., Clément, F., & Harris, P. L. (2004). Trust in testimony: Children's use of true and false statements. *Psychological Science, 15*(10), 694-698. doi:10.1111/j.0956-7976.2004.00742.x
- Kuchirko, Y., Tamis-LeMonda, C. S., Luo, R., & Liang, E. (2015). 'What happened next?': Developmental changes in mothers' questions to children. *Journal of Early Childhood Literacy, 14*68798415598822.
- Kuczaj, S. A. (1977). The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior, 16*(5), 589-600. doi:10.1016/S0022-5371(77)80021-2
- Kurkul, K. E., & Corriveau, K. H. (in press). Question, explanation, follow-up: A mechanism for learning from others? *Child Development*.
- Lieven, E., Salomo, D., & Tomasello, M. (2009). Two-year-old children's production of multiword utterances: A usage-based analysis. *Cognitive Linguistics, 20*(3), 481-507. doi:10.1515/COGL.2009.022
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3 ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- MacWhinney, B., & Snow, C. (1990). The child language data exchange system: An update. *Journal of Child Language, 17*(02), 457-472. doi:10.1017/S0305000900013866
- McCune, L., & Vihman, M. M. (1987). Vocal Motor Schemes. *Papers and Reports in Child Language Development, 26*, 72-79.
- Nelson, K. (1989). *Narratives from the crib*. Cambridge, MA: Harvard University Press.

- Nielsen, M. (2006). Copying actions and copying outcomes: Social learning through the second year. *Developmental Psychology*, 42(3), 555-565. doi:10.1037/0012-1649.42.3.555
- O'Neill, D. K. (1996). Two-year-old children's sensitivity to a parent's knowledge state when making requests. *Child Development*, 659-677.
- Olsen-Fulero, L., & Conforti, J. (1983). Child responsiveness to mother questions of varying type and presentation. *Journal of Child Language*, 10(03), 495-520.
- Peters, A. M. (1990). The role of Imitation in the developing syntax of a blind child. *Text & Talk*, 1(3), 311-312.
- Pew Research Center. (2010). *The new demography of american motherhood*. Retrieved from <http://www.pewsocialtrends.org/2010/05/06/the-new-demography-of-american-motherhood/>
- Post, K. (1994). Negative evidence. In J. Sokolov & C. Snow (Eds.), *Handbook of research in language development using CHILDES* (pp. 132-173). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Redfield, D. L., & Rousseau, E. W. (1981). A meta-analysis of experimental research on teacher questioning behavior. *Review of Educational Research*, 51(2), 237-245. doi:10.3102/00346543051002237
- Rowe, M. L., Coker, D., & Pan, B. A. (2004). A comparison of fathers' and mothers' talk to toddlers in low-income families. *Social Development*, 13(2), 278-291.
- Rowe, M. L., Leech, K. A., & Cabrera, N. (2016). Going beyond input quantity: Wh-questions matter for toddlers' language and cognitive development. *Cognitive Science*.
- Rowland, C. F., & Fletcher, S. L. (2006). The effect of sampling on estimates of lexical specificity and error rates. *Journal of Child Language*, 33(4), 859-877.

doi:10.1017/S0305000906007537

Sachs, J. (1983). Talking about the there and then: The emergence of displaced reference in parent-child discourse. In K. E. Nelson (Ed.), *Children's language* (Vol. 4, pp. 1-28). Hillsdale, NJ: Lawrence Erlbaum Associates.

Sage, K. D., & Baldwin, D. (2010). Social gating and pedagogy: Mechanisms for learning and implications for robotics. *Neural Networks*, 23(8), 1091-1098.
doi:10.1016/j.neunet.2010.09.004

Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge university press.

Sénéchal, M. (1997). The differential effect of storybook reading on preschoolers' acquisition of expressive and receptive vocabulary. *Journal of Child Language*, 24(01), 123-138.

Shafto, P., Goodman, N. D., & Frank, M. C. (2012). Learning from others: The consequences of psychological reasoning for human learning. *Perspectives on Psychological Science*, 7(4), 341-351. doi:10.1177/1745691612448481

Shafto, P., Goodman, N. D., & Griffiths, T. L. (2014). A rational account of pedagogical reasoning: Teaching by, and learning from, examples. *Cognitive Psychology*, 71, 55-89.
doi:10.1016/j.cogpsych.2013.12.004

Shatz, M. (1978). Children's comprehension of their mothers' question-directives. *Journal of Child Language*, 5(01), 39-46.

Shatz, M. (1979). How to do things by asking: Form-function pairings in mothers' questions and their relation to children's responses. *Child Development*, 1093-1099.

Shneidman, L., Gaskins, S., & Woodward, A. (2015). Child-directed teaching and social learning at 18 months of age: evidence from Yucatec Mayan and US infants. *Developmental*

science. doi:10.1111/desc.12318

- Shneidman, L., Gweon, H., Schulz, L. E., & Woodward, A. L. (2016). Learning from others and spontaneous exploration: A cross-cultural investigation. *Child Development, 87*(3), 723-735.
- Siraj-Blatchford, I., & Manni, L. (2008). 'Would you like to tidy up now?' An analysis of adult questioning in the English Foundation Stage. *Early Years, 28*(1), 5-22.
- Snow, C. E. (1972). Mothers' speech to children learning language. *Child Development, 549-565*.
- Snow, C. E. (1983). Literacy and language: Relationships during the preschool years. *Harvard Educational Review, 53*(2), 165-189. doi:10.17763/haer.53.2.t6177w39817w2861
- Snow, C. E., Arlman-Rupp, A., Hassing, Y., Jobse, J., Joosten, J., & Vorster, J. (1976). Mothers' speech in three social classes. *Journal of Psycholinguistic Research, 5*(1), 1-20. doi:10.1007/BF01067944
- Stine, E. L., & Bohannon, J. N. (1983). Imitations, interactions, and language acquisition. *Journal of Child Language, 10*(03), 589-603. doi:10.1017/S0305000900005389
- Suppes, P. (1974). The semantics of children's language. *American Psychologist, 29*(2), 103. doi:10.1037/h0036026
- Theakston, A. L., Lieven, E. V., Pine, J. M., & Rowland, C. F. (2001). The role of performance limitations in the acquisition of verb-argument structure: An alternative account. *Journal of Child Language, 28*(01), 127-152.
- Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.
- Topál, J., Gergely, G., Miklósi, Á., Erdőhegyi, Á., & Csibra, G. (2008). Infants' perseverative search errors are induced by pragmatic misinterpretation. *Science, 321*(5897), 1831-1834.

doi:10.1126/science.1161437

- Van Houten, L. J. (1986). *The role of maternal input in the acquisition process: The communicative strategies of adolescent and older mothers with the language learning children*. Paper presented at the Boston University Conference on Language Development, Boston.
- Von Secker, C. (2002). Effects of inquiry-based teacher practices on science excellence and equity. *The Journal of Educational Research, 95*(3), 151-160. doi:10.2307/27542373
- Vredenburgh, C., Kushnir, T., & Casasola, M. (2014). Pedagogical cues encourage toddlers' transmission of recently demonstrated functions to unfamiliar adults. *Developmental science, n/a-n/a*. doi:10.1111/desc.12233
- Walker, K., & Armstrong, L. (1995). Do mothers and fathers interact differently with their child or is it the situation which matters? *Child: Care, Health and Development, 21*(3), 161-181.
- Warren-Leubecker, A., & Bohannon III, J. N. (1984). Intonation patterns in child-directed speech: Mother-father differences. *Child Development, 1379-1385*. doi:10.2307/1130007
- Weist, R. M., Pawlak, A., & Hoffman, K. (2009). Finiteness systems and lexical aspect in child Polish and English. *Linguistics, 47*(6), 1321-1350. doi:10.1515/LING.2009.046
- Weist, R. M., & Zevenbergen, A. A. (2008). Autobiographical memory and past time reference. *Language Learning and Development, 4*(4), 291-308. doi:10.1080/15475440802293490
- Wells, C. G. (1981). *Learning through interaction: The study of language development*. Cambridge, UK: Cambridge University Press.
- Wilson, B., & Peters, A. M. (1988). What are you cookin' on a hot?: Movement constraints in the speech of a three-year-old blind child. *Language, 64*(2), 249-273.

- Wilson, J., & Henry, A. (1998). Parameter setting within a socially realistic linguistics. *Language in Society*, 27(01), 1-21. doi:10.1017/S0047404500019709
- Wise, K. C., & Okey, J. R. (1983). A meta-analysis of the effects of various science teaching strategies on achievement. *Journal of Research in Science Teaching*, 20(5), 419-435. doi:10.1002/tea.3660200506
- Yu, Y., Landrum, A., Bonawitz, E., & Shafto, P. (under review). Questioning supports effective transmission of knowledge and increased exploratory learning in pre-kindergarden children.

Table 1

Descriptions of Transcripts Used in the Study

Original study	# of mother-child conversations	# of father-child conversations	Era of data collection	Location of data collection	Child's age (month)	Child's gender	Family SES	Sampling and procedural characteristics of the recordings	Average # of parent sentences	Average # of parent questions
(Bohannon III & Marquis, 1977); (Stine & Bohannon, 1983)	1	0	70s	USA	36	M	NS	Conversation between mother and child, with play materials available	213	82
(Braunwald, 1976)	1	0	70s	USA	37	F	NS	Recordings during meal in the researcher's home	20	5
(Demetras, 1986, 1989)	3	4	80s	USA	24-26	4M	NS	20-minute recordings during free play with toys	319.3	150.1
(Demetras, Post, & Snow, 1986; Post, 1994)	2	0	80s	USA	24-31	2F	Working-class	Recordings during free play with toys	776.5	213.5
(Demuth, Culbertson, & Alter, 2006)	6	3	00s	USA	24-27	3M 3F	NS	Recordings during spontaneous interactions at home	479.2	131
(Dickinson & Tabors, 2001)	58	18	80s	USA	43-68	30M 28F	Low-income	Typical mealtime interactions	153.9	51.7
(Feldman & Menn, 2003)	1	1	90s	USA	25	M	NS	10-minute recording in the researcher's home	125	38
(Gleason, 1980)	13	13	70s	USA	32-62	6M 7F	NS	Conversations at dinner	201.2	60.3
(Hayes & Ahrens, 1988)	2	0	80s	USA	36-40	1M 1F	Working-class	Samples of naturalistic speech	93	13.5
(Henry, 1995; J. Wilson & Henry, 1998)	3	1	90s	UK	33-54	2M 1F	NS	Recordings at mealtimes, bedtime, bath time, etc.	341	184
(Howe, 1981)	12	0	70s	UK	24	7M 2F	Specified for each family	Recordings while playing with toys	213	82
(Kuczaj, 1977)	1	1	70s	USA	36	M	NS	30-minute recording in the researcher's home	59	34.5
(Lieven, Salomo, &	1	0	00s	UK	36	M	Middle-	60-minute recording	1063	335

Tomasello, 2009) (MacWhinney, 2000)	0	1	80s	USA	30	M	class NS	during free play Natural record of the researcher's family interactions	58	23
(McCune & Vihman, 1987)	4	0	80s	USA	24	2M 2F	NS	Recordings during free play with toys	225.3	64
(Nelson, 1989)	0	1	80s	USA	24	F	NS	Bedtime conversation	23	7
(Peters, 1990; B. Wilson & Peters, 1988)	0	1	80s	USA	24	M	NS	Conversation between the researcher and his child	251	39
(Rowland & Fletcher, 2006)	1	1	90s	UK	36	F	NS	Recording during everyday play activities	365.5	68
(Sachs, 1983)	1	1	70s	USA	39	F	NS	Conversations in the researcher's home	37	10.5
(Snow, 1983)	1	1	70s	USA	30	M	NS	Conversations in the researcher's home	222.5	50.5
(Suppes, 1974)	1	1	70s	USA	36	F	NS	Naturalistic conversations	369.5	209.5
(Theakston, Lieven, Pine, & Rowland, 2001)	10	2	90s	UK	24-32	5M 5F	NS	30-minute recordings during normal play activities	441.2	160.8
(Van Houten, 1986)	20	0	80s	USA	28	13M 7F	NS	3-minute interactions while the child eats lunch	35	11.9
(Warren-Leubecker & Bohannon III, 1984)	11	10	80s	USA	24-70	6M 5F	Middle- class	Conversation in the child's home, with the child's own toys or books present	153.3	68.7
(Weist, Pawlak, & Hoffman, 2009; Weist & Zevenbergen, 2008)	1	0	00s	USA	38	M	Middle- class	30-minute caregiver-child interactions	256	159
(Wells, 1981)	11	4	70s	UK	26-60	5M 7F	NS	Spontaneous conversation without awareness of being observed	90.5	26
Unpublished research by Julie McMillan	1	0	00s	USA	28	F	NS	Naturalistic conversation	59	26

Note. SES = socioeconomic status; NS = not specified.

Table 2

Coding Scheme for Categories and Subcategories of Questions, and Their Relative Distributions in Mother-child and Father-child Conversations

Category / subcategory	Description	Examples	Proportion in mother-child questions (<i>SD</i>)	Proportion in father-child questions (<i>SD</i>)
Pedagogical	Questioner knows the answer, wants questionee to learn	—	27% (25%)	31% (30%)
Generic	Teaching about kinds of objects or people, general concepts, rules, or scripts	“What’s ‘A’ stand for?”; “What would you say [in this situation]?”	3% (13%)	3% (18%)
Specific	Teaching about a specific object, event, or person	“What’s that?”; “What does this button do?”	24% (24%)	28% (27%)
Information-seeking	Questioner seeks answer from questionee	—	60% (26%)	60% (28%)
Specific	Asking about a specific object, event, or person	“What did you do at school?”	28% (23%)*	35% (24%)*
Check status	Asking about the child’s needs, opinions, or physical/emotional/epistemic status	“Are you hungry?”; “Do you remember?”	24% (22%) [†]	18% (20%) [†]
Clarification	Asking the child to repeat what he/she just said	“You what?”; “Huh?”	7% (10%)	5% (9%)
Permission	Asking for permission	“Can I get you changed?”	1% (4%)	1% (5%)
Rhetorical	Questions not intended to be answered verbally	—	13% (15%)*	9% (11%)*
Commands	Giving commands in a question form	“Why don’t you help clean up?”	8% (12%)*	3% (8%)*
Attention	Raising child’s attention with a question	“Well?”; “Jack?”	5% (8%)	6% (9%)

Note. Asterisks denote significant differences between mother-child and father-child questions, [†] $p < .1$; * $p < .05$; *** $p < .001$.

Table 3

Fixed and Random Effects for Models Predicting the Proportion of Pedagogical Questions

	Model 1				Model 2			
	Estimate	SE	p	VIF	Estimate	SE	p	VIF
Intercept	.13	.15	.41	—	.20	.13	.12	—
Fixed effects (B) on conversation level								
Parent's gender (male - female)	-.004	.041	.93	1.36				
Child's age	-.004	.002	.022*	1.81	-.004	.002	.038*	1.36
Context								
Free play - daily routine	.088	.074	.24	4.53				
Mealtime - daily routine	.087	.087	.33	4.83				
Other parent (present - absent)	.045	.042	.29	1.56				
Sibling (present - absent)	-.028	.041	.50	1.49				
Fixed effects (B) on child level								
Child's gender (male - female)	.045	.032	.16	1.04				
Family SES								
Middle-class - working-class	.224	.088	.015*	2.07	.261	.081	.003**	1.22
Not specified - working-class	.020	.077	.80	1.87	.000	.070	1.00	1.56
Fixed effects (B) on study level								
Time of data collection	.069	.032	.045*	1.22	.063	.028	.038*	1.03
Location (UK - USA)	.005	.073	.95	1.65				
# of children included	-.018	.033	.58					
Sampling (random sample – convenience sample)	.011	.112	.92					
Random effect (σ^2) on study	.009	.007	.21	—	.007	.006	.35	—
-2 restricted log likelihood		44.3				20.2		
BIC		55.1				31.0		

Note. The restricted maximum likelihood (REML) method was used for parameter estimation. *SE* = standard error; VIF = Variance Inflation Factor; SES = socioeconomic status; BIC = Schwarz's Bayesian Information Criterion. * $p < .05$; ** $p < .01$.

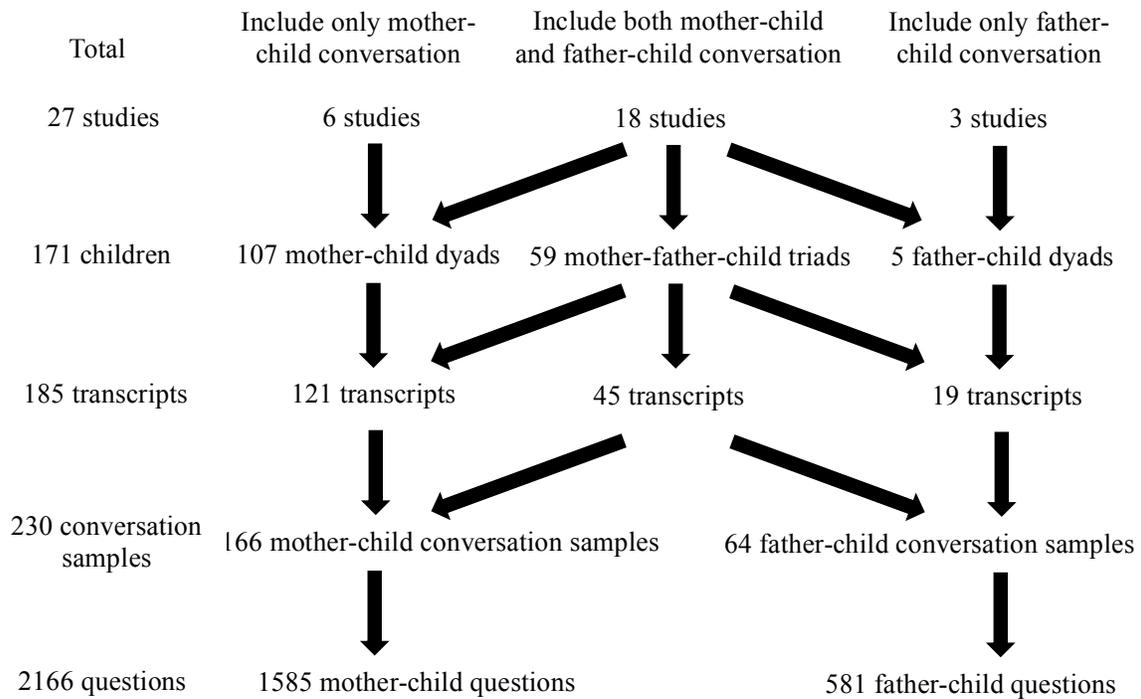


Figure 1. Samples used for analysis. We included 166 mother-child conversation samples and 64 father-child conversation samples from 185 transcripts that represent day-to-day family conversations at home. To ensure independent sampling, only one conversation sample is included for each mother-child and father-child dyad. For each conversation sample, two coders coded the first 10 questions from the parent to the child. If there was less than 10 questions in the sample (this is true for 13% of all samples), all questions were coded. Conversation samples with less than 3 questions were excluded.

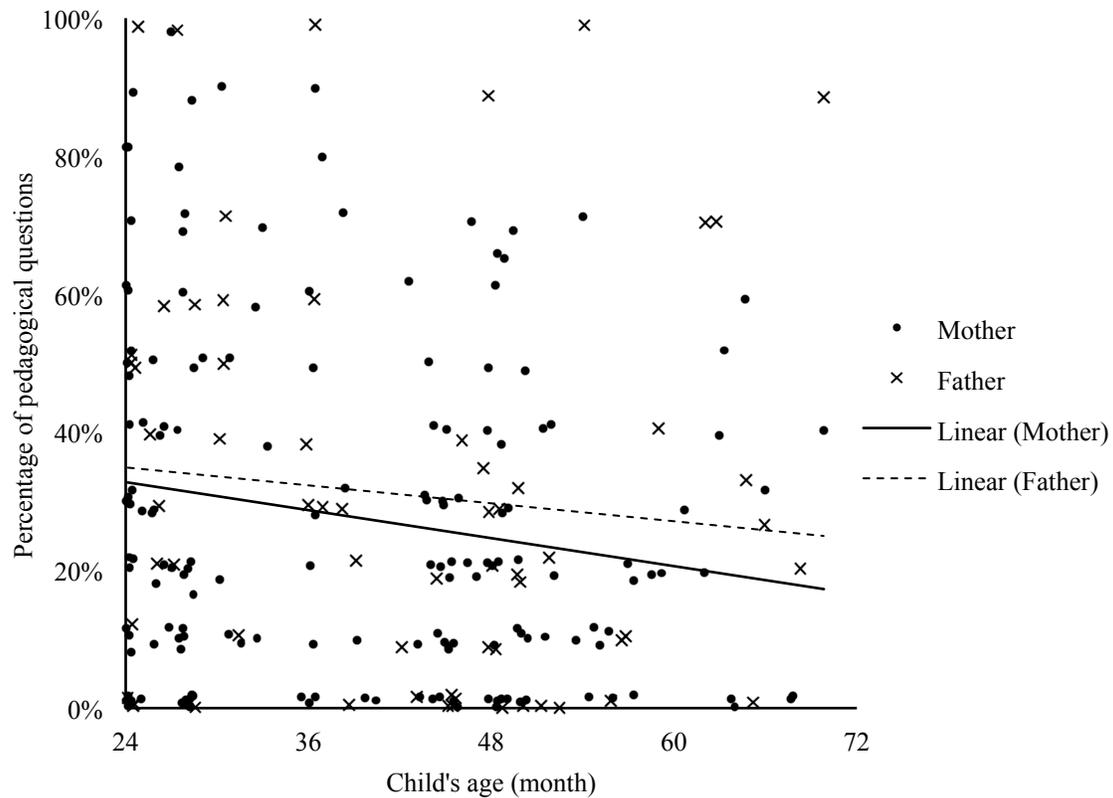


Figure 2. Proportion of parents' pedagogical questions decreased with child's age, $B = -.004, p = .038$. When data were separated by parents' gender, this trend was significant for mothers, $B = -.005, p = .009$, but not for fathers, $B = .001, p = .085$. Possible heteroscedasticity (unequal variabilities in parents' pedagogical questions as a function of child's age) was ruled out by further analyses (details see Supplementary Material). Data points have been jittered along both x axis (within ± 0.5 month) and y axis (within $\pm 2\%$) to improve readability.

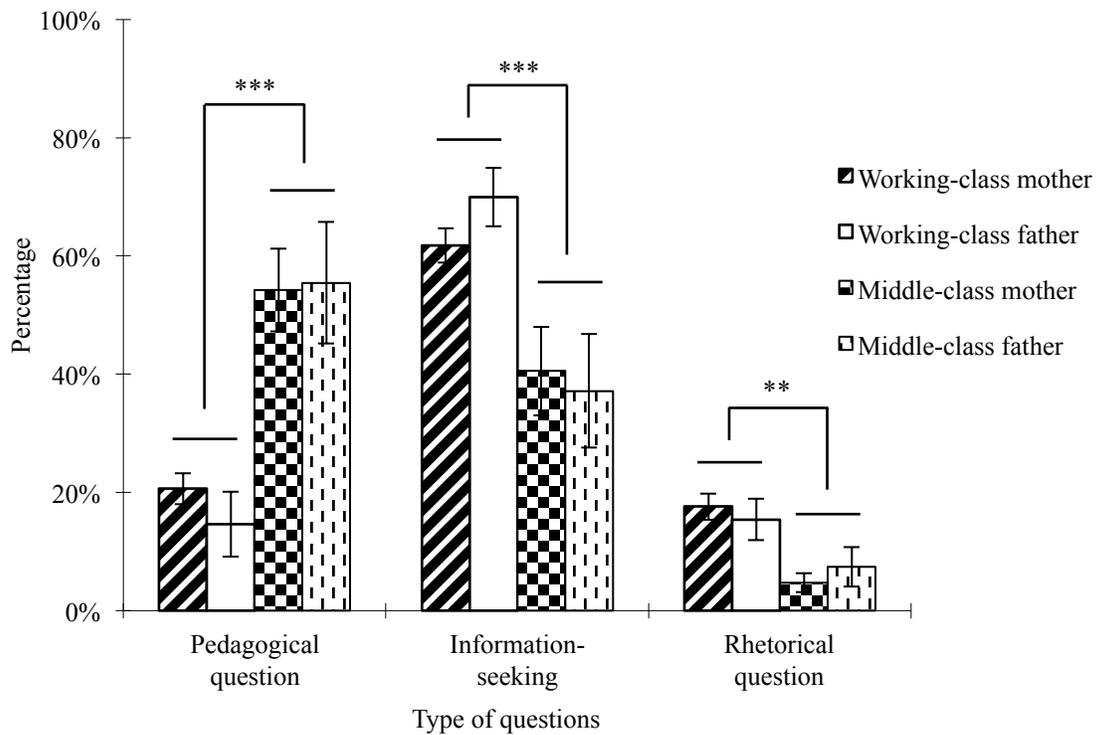


Figure 3. Parents from middle-class families asked more pedagogical questions, fewer information-seeking questions, and fewer rhetorical questions than parents from working-class families, $p_s < .01$. When data were separated by parents' gender, these differences were significant for both mothers and fathers, $p_s < .02$. Further analyses showed that for studies that recruited both working-class and middle-class families, the SES effects on questioning were also observed within studies (details see Supplementary Material).

Error bars denote SE. ** $p < .01$; *** $p < .001$.

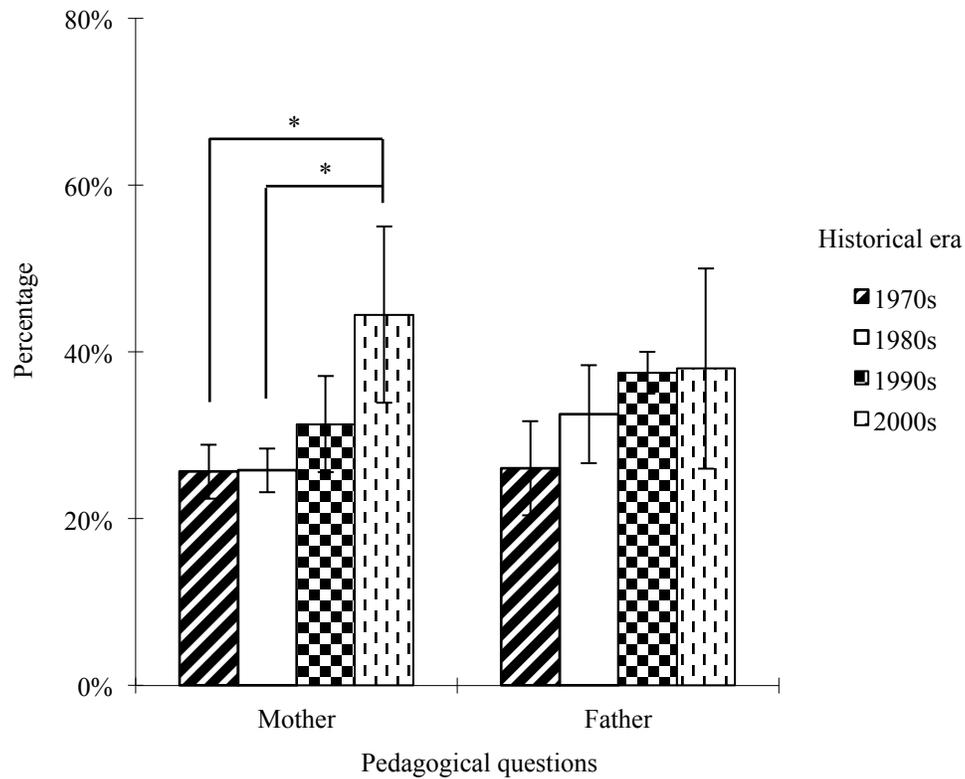


Figure 4. Parents’ pedagogical questions increased with historical era, $B = .063, p = .038$.

When data were separated by parents’ gender, mothers’ pedagogical questions increased significantly with historical era, $B = .054, p = .049$, with mothers from the 2000s asking a higher proportion of pedagogical questions than mothers from the 1970s and 1980s, $ts > 2, ps < .05, ds > 0.7$. The increasing trend of fathers’ pedagogical questions along historical era was not statistically significant, $B = .045, p = .12$. Error bars denote SE. * $p < .05$.

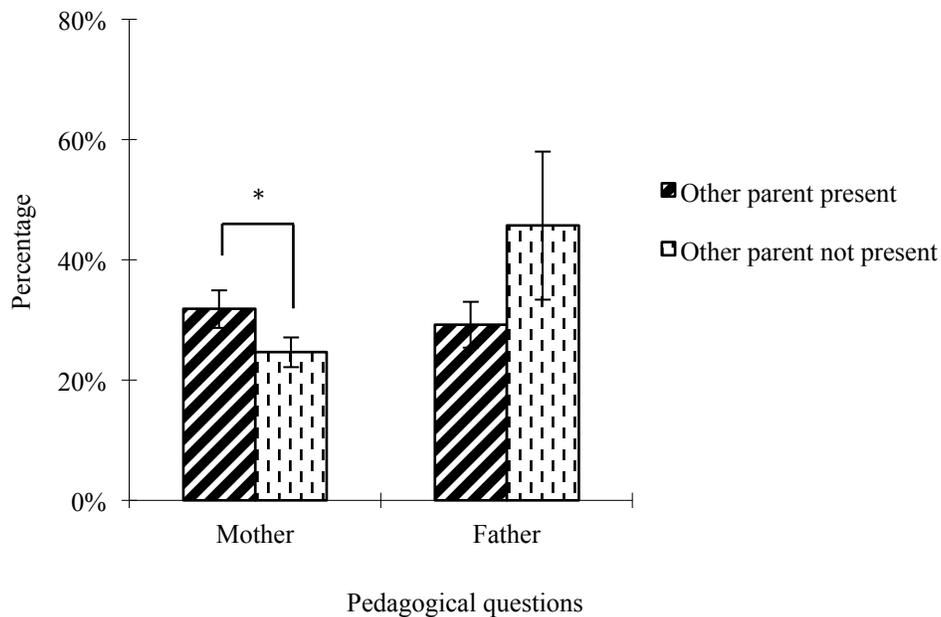


Figure 5. An interaction effect between parents' gender and the presence of other parent was observed for the proportion of pedagogical questions, $F(1, 226) = 4.36, p = .038$. Mothers asked a higher proportion of pedagogical questions when the father was noted as present during the conversation, as compared to when the father was not noted as present, $B = .083, p = .045$. Fathers' pedagogical questions did not differ significantly based on mothers' presence, $B = -.172, p = .17$. This non-significant result may be due to a small but variable sample of father-child conversation when a mother is absent ($n = 19$). Error bars denote SE. * $p < .05$.

Supplementary Material

Detailed rationale for coding and analytical choices

Justification for coding the first 10 questions for each transcript

We put a limit on the maximum questions coded from each conversation to ensure data are balanced across participants, so that no individual parent-child dyad is over-represented in the sample just because they happened to have longer conversations recorded in the database. The choice of making the maximum 10 is because many transcripts have just over 10 questions (8 conversations have exactly 10 questions in them, 30 conversations have between 10 and 20 questions), so we decided that 10 questions is an optimal level to balance between maximizing usable conversations and coding enough questions to be representative for that conversation. We coded the first 10 questions to ensure that the coders are blind to the choice of the sample of questions. This way the coders could not handpick questions that are easy to code or questions fitting certain hypotheses, which could have happened if they were instructed to randomly choose 10 questions.

Estimations for the ambiguity in parent's knowledge states and its impact on our results

When the knowledge state of the questioner is ambiguous based on the question itself, the coders first referred to 1) contexts before and after the question, 2) the linked audio or video clips if available. For questions that remain ambiguous they would make their best guess and mark the question for the first author to review. After coding, the first author has verified all inconsistent coding between coders as well as those marked as ambiguous. His estimate is that ~20% of all questions were ambiguous with regard to the questioner's knowledge state if read without any context: Of these, about half can be determined by the immediate context (~5 lines before or after), and some more can be determined by the broader context or the linked audio or video clip. Therefore, it is estimated that less than 10% of all questions are actually ambiguous. For these questions, it is unlikely that the way coders code would systematically bias with regard to one of the predictors (age, SES, historical era), as the coders were blind to these predictors during coding. The random biases, on the other hand, should have limited impact on the group means because of the large amount of questions being coded.

Training procedure for the coders

To avoid subjective biases, the coders and the first author of the manuscript have gone through a thorough training phase to compare and discuss coding before the coders started formal coding. During the training phase, questions that can be ambiguous were sampled from various scenarios. Coding criteria were developed and recorded based on these scenarios. The scenarios being discussed included parent and child talking about objects or events at hand, past events, future or imaginary events, storytelling, parent disciplining the child, etc. The final coding instructions were written in a detailed step-by-step manner to reduce subjective biases and to ensure replicability (attached below), and all coders followed these instructions during formal coding.

Justification for using multilevel models

Multilevel mixed-effects models are recommended when the data to be analyzed

are aggregated from studies designed for different purposes. Because of the nature of archival data, there is always nonequivalence across studies that are subtle and impossible to disentangle (such as details in the instructions given to participants). The random effect we put on study-level was intended to reduce the impact of these study-specific variations on our analysis: It assumed that each study has some unknown systematic error that influences all transcripts in that study, and these study-specific variances were estimated and removed when the fixed effects were being estimated. This is the same approach adopted in meta-analysis studies to reduce the impact of individual studies on the overall estimation of the true effect.

Additional coding and analyses

Examining the construct validity of pedagogical questions

Additional coding was performed to validate if the questions coded as pedagogical were indeed produced with pedagogical intent. As with many psychological constructs, we cannot directly measure intent. We instead use a proxy to assess validity—having a researcher with experience in the domain recode. This goes beyond inter-rater reliability of the coding scheme itself to begin to examine the construct validity of our measurement according to the theoretical framework.

Our independent coder was not involved in the initial coding, but was familiar with research on informal pedagogy. He was instructed to check a subsample of questions coded as pedagogical ($n = 132$) to see “if the questions can be interpreted as intended to teach, where teaching is broadly defined as any endeavor to help children learn”. He was also instructed to write down what he thought the question was teaching about. Among the 132 questions, 123 (93.2%) were coded as clearly intended to teach, including teaching general concepts, labels, object properties, or conventions (20, 15.2%), teaching about (or help child recall) a specific event, object, or person (78, 59.1%), helping child understand her own needs or status (13, 9.8%), helping child understand someone else’s needs or status (7, 5.3%), or interpreting or correcting child’s use of language (4, 3.0%). The rest 9 questions (6.8%) were ambiguous instances in which the parent may be intended to help the child learn, but the transcript did not provide enough context to accurately judge the parent’s intentions. Importantly, the same 9 questions were also ambiguous with regard to whether the questioner knows the answer of the question. Indeed, out of the 132 questions, we did not find a single question for which the questioner clearly knows the answer but was also clearly not intend to teach (rhetorical questions are in a separate category and are not discussed here). Part of the reason, we think, is that just like children, our coder also often relies on the questioner’s knowledge state to infer the intention behind the question. For example, “What did you do at school today?” could be helping children remembering and expressing a past event if the parent knew what happened at school (which would make it a pedagogical question), or it could be trying to learn the child’s school experience if the parent did not know what happened (which would make it an information seeking question). We take these findings as evidence that at least from the learner and coder’s perspective, knowledge state and intention are often closely coupled in the case of pedagogical questions.

Cross-validation of multilevel model

We have used cross-validation to examine the robustness of Model 2. We randomly separated our sample into two equal halves, constructed a model from the first half, and then tested it on the second half. Results showed that the same trends observed for the whole sample (the effects of child's age, SES and historical time) were also observed in the two halves, although some of the effects were marginally significant because of the decreased statistical power. For the first half of data, the model showed a negative trend of child's age, $B = -0.0374$, $p = .063$, a difference between middle-class and working class families, $B = 0.224$, $p = .053$, and an increasing trend through historical time, $B = 0.0686$, $p = .113$. When we apply this model to the second half of data, the model fit was good, -2 restricted log likelihood = 1.10, BIC = 10.50, and the same trends showed up for child's age ($B = -0.0321$, $p = .158$), SES ($B = 0.216$, $p = .055$), and historical time ($B = 0.0772$, $p = .064$).

Logistic regression on the question level

Because all questions were coded as categorical variables, an alternative way to analyze our data is to treat question as a binary dependent variable (e.g., pedagogical versus not pedagogical), and apply logistic regression. We performed this alternative analysis to further validate Model 2, and the results were consistent to the linear model.

To apply this alternative model, we first reconstructed our data so that each data point represented a question rather than a parent-child dyad, and then we applied logistic regression to the restructured data. This was done by using Generalized Linear Mixed Models in SPSS, and using a binary logistic regression as the link function for the dependent variable. Because now the unit of analysis is questions, we also added a random effect on parent-child dyad (in addition to the random effect on study). The results were similar to the original analysis: There were significant effects of child's age ($B = .025$, $p < .001$), SES ($B = .729$, $p = .02$), and historical era ($B = .551$, $p < .001$).

Test for heteroskedasticity for the age effects

Figure 2 in the main manuscript has shown a possible trend of decreasing variance in pedagogical questions along age, which can be a violation of the presumptions for our linear regression model. We used the White test to examine whether the heteroskedasticity (unequal residual along predictor) is significant for the regression model. The White test turned out to be non-significant both when mothers and fathers' data were combined (LM = 3.45, $p = .18$) and when mothers and fathers' data were separate (mother: LM = 5.48, $p = .064$; father: LM = 1.66, $p = .44$). Therefore, based on standard practice, we did not adjust the model by using heteroskedasticity-consistent standard errors.

Test for within-study SES effects

One challenge of analyzing archival data aggregated from multiple studies is the potential heterogeneity in the sampling and procedure of the original studies. Though we made our best effort to document and control for possible sampling and procedural differences, there can still be subtle difference that are not documented in the database. Therefore, we examined if one of the key factors found in the study, and one that may be subject to the influence of subtle differences in study design—the SES differences—is robust within a study. There is one study (Howe, 1981) which compared 6 middle-class and 6 working-class families and recorded mother-child conversations through the same

procedure. On average the middle-class mothers asked almost twice as many PQs as the working-class mothers (38.3% vs. 20%), though the result was not statistically significant due to small sample. This suggests that SES differences do not just reflect differences in study design.

Step-by-step coding instructions

Overall process

1. Have the coding sheet open and find your tag.
2. Go to [CHILDES database](#), find the file.
3. Get the total number of mother's or father's utterances by running `mdu +t*MOT filename.cha` or `mdu +t*FAT filename.cha` (select "mdu" and put `t*MOT / FAT filename.cha` in the box to the right, replace "filename" with the actual file name). Plug the number in the "Common codes" tag.
4. Get the total number of mother's or father's questions by running `combu +s? +t*MOT filename.cha` or `combu +s? +t*FAT filename.cha`, the total number is in the end of the output. Plug the number in the "Common codes" tag.
5. We code the first 10 questions for each file (or all questions if there are less than 10 questions). Plug the line numbers in the "Common codes" tag.
6. Tick off all depend tiers
7. Highlight all question marks (Command-F, 2"?)")
8. Code Function, Answer, and Follow up as detailed below
9. Code context based on the script you coded (some documents have "situations" listed on the top, these refer to the whole conversation, and can be different from the part of conversation you coded). Plug the code in the "Common codes" tag.

We code all sentences that:

1. Has a question mark in the end.
2. Is from a mother or father to the target child. (Don't code a question asked by the child, or a question from the parent to another parent/sibling, etc.)
3. If there is a question without a question mark or any question marks in brackets "[?]", just ignore it. *
4. Skip if a question contains missing words "xxx" or "<"

Each question should get a code of 4 letters, and some questions would also get a number. Each question needs to be coded along the following dimensions:

Function: First determine if the questioner knows the answer of the question (P), doesn't know the answer (I), or if it is a rhetorical question for which no answer is expected (R).

When the questioner knows the answer of the question (P), assign one of following:

PG – Teaching about generic knowledge--knowledge about kinds or concepts ("Do birds have wings?", "what is 1+1?", "what's a@l stand for?")

PS - Teaching about specific knowledge, such as an object, an activity/event, a person, etc., ("Is this tree green or

red?", "Who came for your birthday party?", "You putting the doggie in a chair?", "That car doesn't fit?")**

For confirmation questions - those when the parent repeated what the child said in a question form - we no longer have a separate category for them. Those questions need to be coded into other categories based on whether the parent knows the answer of the question and what the question was for. They might be IC, PS, IN, IS, or others.

When the questioner does not know the answer of the question (I):

IC - Clarification, asking the child to repeat what he/she just said because it wasn't clearly heard ("huh?" "what?"), or to clarify what the child meant. Not when correcting children's language or pronunciation (which would be PS)

IS - Ask about an object, an activity/event, or a person ("where's your turtle?", "how many games did you play?", "What are you doing?")**

IN - Check the child's needs or physical/emotional status ("Do you have to go potty?", "Do you want ...?", "What do you want to do?"); Check the child's memory, thinking, or opinion ("Did you remember that?", "Did you know that?", "Have you thought about that?" "Did you dream about something?"); Or offering help ("Do you want me to help?")

IP - Asking for permission ("Can I sit here?", but "Can I have that?" is RC because it requires the child to act)

When no verbal answer is expected, rhetorical questions (R)

RC - Command for children to do something, except for talking ("Can you help me...?", "Why don't you...?")

RE - Elicit children's attention ("See?", "Well?", "What did I say?", "Jack?")

Answer: Code + number

AS - Self answer: if the questioner followed the question with an answer

AQ - Add question: if the questioner followed the question with another question

AN - No answer: if the next person who speaks is not the child, and is not AS/AQ

AI - Irrelevant answer: if the next person who speaks is the child, but what he/she said is irrelevant to the question

AY - Relevant answers + number of words in that sentence

AR - Repeating the question or part of the question + number of words in that sentence

Notes: * When a sentence omits the beginning ("See that?") and can be interpreted as either a question ("Did you see that?") or a statement ("You see that?"), code it as a question.

** A same question can have different codes under different contexts, depending on

whether the questioner knows the answer. If it is hard to judge whether the questioner knows the answer from the script, listen to the audio/video. If no audio/video is linked, make a best guess and put a question mark after your code.