Pedagogical Questions in Parent–Child Conversations

Yue Yu, Elizabeth Bonawitz, and Patrick Shafto
Rutgers University–Newark

Questioning is a core component of formal pedagogy. Parents commonly question children, but do they use questions to teach? This article defines “pedagogical questions” as questions for which the questioner already knows the answer and intended to help the questionee learn. Transcripts of parent–child conversations were collected from the CHILDES database to examine the frequency and distribution of pedagogical questions. Analysis of 2,166 questions from 166 mother–child dyads and 64 father–child dyads (child’s age between 2 and 6 years) showed 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 toward understanding the role of parent–child questions in facilitating children’s learning.

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 about 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 5 months old (Bornstein et al., 1992). Questions 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 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

This study is supported by NSF CAREER, DRL-1149116 to Patrick Shafto, NSF Awards SMA-1640816 and ECR-1660885 to Elizabeth Bonawitz and Patrick Shafto, and a Jacob Foundation Research Fellowship to Elizabeth Bonawitz. 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 Psychology, Rutgers University–Newark, Newark, NJ 14850. Electronic mail may be sent to yue.yu@rutgers.edu.

© 2017 The Authors
Child Development © 2017 Society for Research in Child Development, Inc.
All rights reserved. 0009-3920/2017/xxxx-xxxx
DOI: 10.1111/cdev.12850
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 questions 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 article 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, and so forth (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 facilitate 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; Vredenburgh, Kushnir, & Casasola, 2014).

Like pedagogical instruction, pedagogical questions imply a learning opportunity 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: Whereas 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 affords 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.
Pedagogical Questions in CHILDES

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 intended 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 parent’s 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). The transcripts were obtained from the CHILDES database between October 2015 and February 2017.
Coding Procedures

Step-by-step coding instructions and the training process of the coders are detailed in the Supporting Information. 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 (a) ended with a question mark, (b) were asked by the parent, (c) were directed toward the target child, and (d) did not contain missing words. Of the 230 conversation samples, 30 contained < 10 such questions (with a minimum of three 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 (a) contexts before and after the question, (b) the linked audio or video clips if they were available. Because 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 the Supporting Information, 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 interrater 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 $k = .830$; when all eight subcategories were considered, Cohen’s $k = .832$. Inconsistent codes were reviewed and resolved by a third coder.
Table 1  
*Descriptions of Transcripts Used in the Study*

<table>
<thead>
<tr>
<th>Original study</th>
<th>No. of mother-child conversations</th>
<th>No. of father-child conversations</th>
<th>Era of data collection</th>
<th>Location of data collection</th>
<th>Child’s age (months)</th>
<th>Child’s gender</th>
<th>Family SES</th>
<th>Sampling and procedural characteristics of the recordings</th>
<th>Average no. of parent sentences</th>
<th>Average no. of parent questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bohannon and Marquis (1977) and Stine and Bohannon (1983)</td>
<td>1</td>
<td>0</td>
<td>1970s USA</td>
<td>36</td>
<td>M</td>
<td>NS</td>
<td>Conversation between mother and child, with play materials available</td>
<td>213</td>
<td>82</td>
<td></td>
</tr>
<tr>
<td>Braunwald (1971)</td>
<td>1</td>
<td>0</td>
<td>1970s USA</td>
<td>37</td>
<td>F</td>
<td>NS</td>
<td>Recordings during meal in the researcher’s home</td>
<td>20</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Demetras (1986, 1989)</td>
<td>3</td>
<td>0</td>
<td>1980s USA</td>
<td>24–26</td>
<td>4M</td>
<td>NS</td>
<td>20-min recordings during free play with toys</td>
<td>319.3</td>
<td>150.1</td>
<td></td>
</tr>
<tr>
<td>Demetras, Post, and Snow (1986) and Post (1994)</td>
<td>0</td>
<td>0</td>
<td>1980s USA</td>
<td>24–31</td>
<td>2F</td>
<td>Working-class</td>
<td>Recordings during free play with toys</td>
<td>776.5</td>
<td>213.5</td>
<td></td>
</tr>
<tr>
<td>Demuth, Culbertson, and Alter (2006)</td>
<td>6</td>
<td>3</td>
<td>2000s USA</td>
<td>24–27</td>
<td>3M, 3F</td>
<td>NS</td>
<td>Recordings during spontaneous interactions at home</td>
<td>479.2</td>
<td>131</td>
<td></td>
</tr>
<tr>
<td>Dickinson and Tabor (2001)</td>
<td>58</td>
<td>18</td>
<td>1980s USA</td>
<td>43–68</td>
<td>30M, 28F</td>
<td>Low-income</td>
<td>Typical mealtime interactions</td>
<td>153.9</td>
<td>51.7</td>
<td></td>
</tr>
<tr>
<td>Feldman and Mens (2003)</td>
<td>1</td>
<td>1</td>
<td>1990s USA</td>
<td>25</td>
<td>M</td>
<td>NS</td>
<td>10-min recording in the researcher’s home</td>
<td>125</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>Gleason (1980)</td>
<td>13</td>
<td>13</td>
<td>1970s USA</td>
<td>32–62</td>
<td>6M, 7F</td>
<td>NS</td>
<td>Conversations at dinner</td>
<td>201.2</td>
<td>60.3</td>
<td></td>
</tr>
<tr>
<td>Hayes and Ahrens (1988)</td>
<td>2</td>
<td>0</td>
<td>1980s USA</td>
<td>36–40</td>
<td>1M, 1F</td>
<td>Working-class</td>
<td>Samples of naturalistic speech</td>
<td>93</td>
<td>13.5</td>
<td></td>
</tr>
<tr>
<td>Henry (1995) and J. Wilson and Henry (1998)</td>
<td>3</td>
<td>1</td>
<td>1990s UK</td>
<td>33–54</td>
<td>2M, 1F</td>
<td>NS</td>
<td>Recordings at mealtimes, bedtime, bath time, etc.</td>
<td>341</td>
<td>184</td>
<td></td>
</tr>
<tr>
<td>Howe (1981)</td>
<td>12</td>
<td>0</td>
<td>1970s UK</td>
<td>24</td>
<td>7M, 2F</td>
<td>Specified for each family</td>
<td>Recordings while playing with toys</td>
<td>213</td>
<td>82</td>
<td></td>
</tr>
<tr>
<td>Kuczaj (1977)</td>
<td>1</td>
<td>1</td>
<td>1970s USA</td>
<td>36</td>
<td>M</td>
<td>NS</td>
<td>30-min recording in the researcher’s home</td>
<td>59</td>
<td>34.5</td>
<td></td>
</tr>
<tr>
<td>Lieven, Salomo, and Tomasello (2009)</td>
<td>1</td>
<td>0</td>
<td>2000s UK</td>
<td>36</td>
<td>M</td>
<td>Middle-class</td>
<td>60-min recording during free play</td>
<td>1063</td>
<td>335</td>
<td></td>
</tr>
<tr>
<td>MacWhinney (2000)</td>
<td>0</td>
<td>1</td>
<td>1980s USA</td>
<td>30</td>
<td>M</td>
<td>NS</td>
<td>Natural record of the researcher’s family interactions</td>
<td>58</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>Original study</td>
<td>No. of mother-child conversations</td>
<td>No. of father-child conversations</td>
<td>Era of data collection</td>
<td>Location of data collection</td>
<td>Child’s age (months)</td>
<td>Child’s gender</td>
<td>Family SES</td>
<td>Sampling and procedural characteristics of the recordings</td>
<td>Average no. of parent sentences</td>
<td>Average no. of parent questions</td>
</tr>
<tr>
<td>--------------------------------------------</td>
<td>-----------------------------------</td>
<td>-----------------------------------</td>
<td>------------------------</td>
<td>----------------------------</td>
<td>----------------------</td>
<td>----------------</td>
<td>------------</td>
<td>--------------------------------------------------------</td>
<td>-----------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>McCune and Vihman (1987)</td>
<td>4</td>
<td>0</td>
<td>1980s USA</td>
<td>24</td>
<td>2M, 2F</td>
<td>NS</td>
<td></td>
<td>Recordings during free play with toys</td>
<td>225.3</td>
<td>64</td>
</tr>
<tr>
<td>Nelson (1989)</td>
<td>0</td>
<td>1</td>
<td>1980s USA</td>
<td>24</td>
<td>F</td>
<td>NS</td>
<td></td>
<td>Bedtime conversation</td>
<td>23</td>
<td>7</td>
</tr>
<tr>
<td>Peters (1990) and B. Wilson and Peters (1988)</td>
<td>0</td>
<td>1</td>
<td>1980s USA</td>
<td>24</td>
<td>M</td>
<td>NS</td>
<td></td>
<td>Conversation between the researcher and his child</td>
<td>251</td>
<td>39</td>
</tr>
<tr>
<td>Rowland and Fletcher (2006)</td>
<td>1</td>
<td>1</td>
<td>1990s UK</td>
<td>36</td>
<td>F</td>
<td>NS</td>
<td></td>
<td>Recording during everyday play activities</td>
<td>365.5</td>
<td>68</td>
</tr>
<tr>
<td>Sachs (1983)</td>
<td>1</td>
<td>1</td>
<td>1970s USA</td>
<td>39</td>
<td>F</td>
<td>NS</td>
<td></td>
<td>Conversations in the researcher’s home</td>
<td>37</td>
<td>10.5</td>
</tr>
<tr>
<td>Snow (1983)</td>
<td>1</td>
<td>1</td>
<td>1970s USA</td>
<td>30</td>
<td>M</td>
<td>NS</td>
<td></td>
<td>Conversations in the researcher’s home</td>
<td>222.5</td>
<td>50.5</td>
</tr>
<tr>
<td>Suppes (1974)</td>
<td>1</td>
<td>1</td>
<td>1970s USA</td>
<td>36</td>
<td>F</td>
<td>NS</td>
<td></td>
<td>Naturalistic conversations</td>
<td>369.5</td>
<td>209.5</td>
</tr>
<tr>
<td>Theakston, Lieven, Pine, and Rowland (2001)</td>
<td>10</td>
<td>2</td>
<td>1990s UK</td>
<td>24–32</td>
<td>5M, 5F</td>
<td>NS</td>
<td></td>
<td>30-min recordings during normal play activities</td>
<td>441.2</td>
<td>160.8</td>
</tr>
<tr>
<td>Van Houten (1986)</td>
<td>20</td>
<td>0</td>
<td>1980s USA</td>
<td>28</td>
<td>13M, 7F</td>
<td>NS</td>
<td></td>
<td>3-min interactions while the child eats lunch</td>
<td>35</td>
<td>11.9</td>
</tr>
<tr>
<td>Warren-Leubecker and Bohannon (1984)</td>
<td>11</td>
<td>10</td>
<td>1980s USA</td>
<td>24–70</td>
<td>6M, 5F</td>
<td>Middle-class</td>
<td></td>
<td>Conversation in the child’s home, with the child’s own toys or books present</td>
<td>153.3</td>
<td>68.7</td>
</tr>
<tr>
<td>Weist, Pawlak, and Hoffman (2009) and Weist and Zevenbergen (2008)</td>
<td>1</td>
<td>0</td>
<td>2000s USA</td>
<td>38</td>
<td>M</td>
<td>Middle-class</td>
<td></td>
<td>30-min caregiver–child interactions</td>
<td>256</td>
<td>159</td>
</tr>
<tr>
<td>Wells (1981)</td>
<td>11</td>
<td>4</td>
<td>1970s UK</td>
<td>26–60</td>
<td>5M, 7F</td>
<td>NS</td>
<td></td>
<td>Spontaneous conversation without awareness of being observed</td>
<td>90.5</td>
<td>26</td>
</tr>
<tr>
<td>Unpublished research by Julie McMillan</td>
<td>1</td>
<td>0</td>
<td>2000s USA</td>
<td>28</td>
<td>F</td>
<td>NS</td>
<td></td>
<td>Naturalistic conversation</td>
<td>59</td>
<td>26</td>
</tr>
</tbody>
</table>

*Note. SES = socioeconomic status; NS = not specified.*
To ensure that questions for which parents knew the answer 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 remaining nine 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 Supporting Information.

For each question, the two initial coders also decided whether the child responded to the question (repeated the question or provided an answer), the parent followed up the question (repeated the question, provided an answer, or added relevant information), or there was no direct response to the question (a different person spoke or the topic was changed). Interrater 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 (a) over meal (lunch or dinner), (b) during free play (e.g., playing with toys, book reading, spontaneous play or talk), or (c) during daily routines (e.g., taking a bath, preparing to leave).

### Data Analysis

All data were entered and analyzed in IBM SPSS 22.0, Armonk, NY, United States. 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

---

Table 2

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

Asterisks denote significant differences between mother–child and father–child questions. \(^* p < .1. * p < .05. *** p < .001.\)
the Supporting Information. 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. 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 toward children. We then reduced Model 1 by excluding all nonsignificant 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 (for details, see the Supporting Information).

Model 2 showed that the proportion of pedagogical questions decreased with children’s age (Figure 2), $B = -0.004, p = .038$. Parents asked a larger proportion of pedagogical questions to toddlers (24–36 months; $n = 72, M = 65$%) than to preschoolers (36–72 months; $n = 94, M = 25$%), $t(179.4) = 2.43, p = .016, d = .33$. When we separated subcategories of generic and specific pedagogical questions, the age trend was significant for the generic questions, $B = -0.004, p < .001$, 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

| Table 3 | Fixed and Random Effects for Models Predicting the Proportion of Pedagogical Questions |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                | 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) | $-0.004$ | .041  | .93   | 1.36 | $-0.004$ | .002  | .022* | 1.81 |
| Child’s age    | $-0.004$ | .002  | .022* | 1.81 | $-0.004$ | .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) | $-0.028$ | .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 | .202     | .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 |           |       |       |       |
| No. of children included | $-0.018$ | .033  | .58   |       |           |       |       |       |
| Sampling (random sample–convenience sample) | .011     | .112  | .92   |       |           |       |       |       |
| Random effect ($\sigma^2$) 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 method was used for parameter estimation. SE = standard error; VIF = variance inflation factor; SES = socioeconomic status; BIC = Bayesian information criterion. *$p < .05$. **$p < .01$. 
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 ($n = 64$, $M = 26\%$), $t(76) = 2.30$, $p = .024$, $d = .68$. 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), $B < -.110$, $ps < .066$. 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, $B < .017$, $ps > .29$. 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 were 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
was 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 > .7. Fathers’ pedagogical questions also showed an increasing trend with historical era (Figure 4), although the trend was nonsignificant, 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 was 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
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, is higher in parents’ questions toward 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 can be sensitive to others’ pedagogical intent and knowledge states (Csibra & Gergely, 2009; O’Neill, 1996), which raises the possibility that pedagogical questions can 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 1980s. 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 are 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 U.S.A. and U.K., 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 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


Demetras, M., Post, K. N., & Snow, C. E. (1986). Feedback to first language learners: The role of repetitions and


**Supporting Information**

Additional supporting information may be found in the online version of this article at the publisher’s website:

**Data S1.** Additional information for method and results