Reconceptualizing Children’s Suggestibility: Bidirectional and Temporal Properties

Livia L. Gilstrap  
University of Colorado  

Stephen J. Ceci  
Cornell University

Forty-one children (3 to 7 years) were exposed to a staged event and later interviewed by 1 of 41 professional interviewers. All interviews were coded with a detailed, mutually exclusive, and exhaustive coding scheme capturing adult behaviors (leading questions vs. neutral) and child behaviors (acquiescence vs. denial) in a temporally organized manner. Overall, interviewers’ use of leading questions did not result in increased acquiescence as previously found. However, one specific type of leading question (i.e., inaccurate misleading) was followed by acquiescence. Lagged sequential analyses showed that it was possible to predict directly from child-to-child behavior, effectively skipping the intervening adult behavior. This result raises questions about the current conceptualization that suggestibility is driven by adult behaviors.

Studies of children’s suggestibility were published as early as the late 19th century (e.g., Small, 1896), but it was not until the last quarter of the 20th century that studies involving preschool children’s suggestibility appeared (see Ceci & Bruck, 1993, for a review). Since then, however, there has been an outpouring of research on young children’s susceptibility to interviewers’ suggestions (Ceci & Bruck, 1993; Qin, Quas, Redlich, & Goodman, 1997). From this research it is clear that children’s autobiographical memory can be highly accurate, but that their reports can be distorted when they are obtained under suggestive interviewing conditions. Factors such as question repetition (Poole & White, 1991, Poole, 1993), use of yes–no questioning (Brady, Poole, Warren, & Jones, 1999; Walker, Lunning, & Eilts, 1996), misleading questions (e.g., Ornstein, Larus, & Clubb, 1991), plausible and implausible suggestions (Pezdek, Finger, & Hodges, 1997; Pezdek & Hodges, 1999), stereotyping (Leichtman & Ceci, 1995), and invoking peer conformity (Scullin, Kanaya, & Ceci, 2002) have been associated with errors in children’s reports to adult interviewers. Moreover, younger children appear more vulnerable to the deleterious effects of an interviewer’s misleading suggestions than older children (Ceci & Friedman, 2000).

This research has been important for both theoretical and practical reasons. Regarding the former, the issue of childhood suggestibility has long occupied a prominent position in theoretical debates because of its centrality to the issue of whether suggestions overwrite memories or whether they merely alter report accuracy as a result of acquiescence to social pressure. The traditional definition of suggestibility, which is “the extent to which individuals come to accept and subsequently incorporate post-event information into their memory recollections,” implies that suggestions wreak their havoc by altering the underlying memory (Gudjonsson, 1986). However, there are many instances in which suggestibility reduces report accuracy not through actual changes in memory but merely because children may acquiesce to false suggestions made to them by interviewers while their original event memories remain unaltered (e.g., McCloskey & Zaragoza, 1985).

The practical importance of suggestibility stems from a series of high-profile acquittals, convictions, and reversals of convictions in child sexual abuse cases, based in part on claims of suggestive interviews (see Ceci & Bruck, 1995, for case descriptions). Today, research on children’s suggestibility informs curricula for training interviewers, preparing witnesses, and educating judges (Poole & Lamb, 1998; Saywitz, 1995). Both theoretical and pragmatic forces have converged
to produce a large corpus of data on the suggestibility of children’s reports to adult interviewers.

The research described thus far has focused on the influence of adults on children’s accurate recall and children’s suggestibility. Notwithstanding its practical and theoretical significance, however, the research on suggestibility has been predicated on a paradigm that is limited—namely, the analysis of structured interviews. To assess adult to child influences, most studies of suggestibility have used a classic memory paradigm in which an event is followed by a series of structured recall questions. Children’s suggestibility is then defined as the number of times children acquiesce to the suggestions. The use of structured interviews results in two major limitations.

First, it is unclear whether the current understanding of adult-to-child influences, based on structured interview assessments, would hold true in unstructured interviews. Leading questions, for example, may have different effects if the interviewer controls when they are asked and to which children they are asked. Furthermore, children may respond differently in a structured interview than in an unstructured interview. Finally, it is important to note that unstructured interviews (e.g., in child sexual abuse cases) are the case to which researchers are trying to generalize. Although several thorough examinations have been performed on unstructured interviews (e.g., see work from Lamb, Sternberg, & Esplin, 1994; Lamb et al., 2002, who have provided extensive reports of the proportion of Israeli and American interviewers’ use of various types of suggestive utterances and the amount of details elicited), these studies have not examined links between suggestive questions and acquiescence. Although field interviewers are constrained broadly by their training (e.g., begin by acclimation, proceed to free narrative, follow with probing, etc.), it is unclear whether current suggestibility findings based on structured, constrained interviews would hold true if field interviewers conducted unstructured, relatively unconstrained interviews. This is in large part because researchers in the field have not used a methodology that makes a fine-grained analysis of an ever-changing conversational stream possible.

Second, missing from a structured interview approach is the influence of the child on the interviewer. Unknown is the extent to which children affect interviewers and children’s previous responses influence their subsequent behaviors, either directly or through their influence on the interviewer. These direct and indirect effects of the child could render the interviewer’s reactions nondeterminative or at best mediators of the child’s ultimate response.

There is some evidence that some child individual differences, for example, affect children’s interview behavior through their influence on adult interview behavior (Gilstrap & Papierno, in press). What has not been possible within this traditional memory paradigm is a more complex structuring and analysis of interviews that would allow a contrast between the effects of the adult and the effects of the child to determine which effects are primary (determinative) and which are merely epiphenomenal. Thus, in addition to the corpus of studies showing that an interviewer’s behavior appears to alter a child’s response, it would be helpful to have an analytic structure that permits the examination of previously unlooked for child-to-adult influences and their interaction with classically studied adult-to-child influences. Such a framework would allow researchers to examine not only the bidirectional influences of the child’s response on the interviewer’s subsequent behavior and vice versa but would also permit researchers to examine the possibility that a child’s denial or acquiescence is the product of his or her own earlier denial or acquiescence rather than the interviewer’s intermediate questioning. It has not been feasible to determine whether the relationship between an interviewer’s behavior and a child’s reply is influenced by the prior statement of the child.

The major reason that such a determination has not been feasible is that most studies of suggestibility have used interviews in which the interviewers’ behaviors are largely predetermined and do not vary as a function of the child’s responses. Before the conversation begins, a script is laid out in which the interviewer’s responses either are the same, regardless of the child’s responses to the recall prompts, or are prespecified according to a decision tree (e.g., If the child replies X, then ask Y). Thus, in structured interviews, interviewers are instructed about what to ask the child, when to ask it, and how to ask it, and the children’s responses are subsequently analyzed in terms of discrete units (e.g., number of acquiescences in response to yes – no questions). This type of experimental design, albeit an important source of information about the potentially deleterious effects of suggestive questions, cannot address the possibility that in unstructured interviews the adult tailors his or her questions to the type of responses the child has already provided, and differentiates his or her questioning as a function of the child’s perceived competencies, intransigence, and personality. Thus, the structured experimental situation controls half of the dyad instead of allowing the conversation to be progressively and jointly determined, and constantly reevaluated. If the adult’s behavior is controlled,
researchers cannot look for influences of the child on the adult.

Only a sequential analysis of the conversation stream permits a simultaneous examination of child-to-adult as well as adult-to-child influences; interrogative suggestibility is by definition an interactive, temporally contingent process. Thus, to study the transmission of suggestions, it is important to study the temporal relationship between interviewer suggestions and child responses in a realistically unstructured conversational stream. This analytic technique allows researchers to go beyond hypothesis-confirming analyses such as “Does suggestive behavior X result in response Y?” to undertake hypothesis-generating analyses such as “What types of interviewer behaviors precede inaccurate responses by the child?”

However, studying unstructured conversation poses its own challenges. Most important, it requires statistical and methodological tools that can extract relationships between adult and child behavior from a nonexperimentally controlled stream of conversation. Fortunately, such techniques have been developed and used with great success by dyad researchers, particularly developmental psychologists (see Bakeman & Gottman, 1997; Bakeman & Quera, 1995, for an excellent introduction to observational coding and sequential analysis).

Sequential analysis is a process-oriented method of data analysis that systematically divides behavior streams into units and assigns meaningful values to each. The data stream can be divided into speech turns: adult–child–adult–child, and each turn can be given a series of codes:

Leading question by interviewer – acquiescence by child – neutral question by interviewer. Sequential analysis is specifically designed to answer the question, “Given behavior code Y, is behavior code X more likely to follow than would be expected by chance?” This method of analysis is termed sequential because it is predicated on the temporal relationship between behaviors. In the current example, one might inquire about adult-to-child influences by asking, “Given that the adult asks a leading question, will the child be more likely than chance to acquiesce?” or about child-to-adult influences by asking, “Given that the child has not acquiesced, will the interviewer be more likely than chance to ask a leading question?” To answer both example questions, transitional probabilities would be computed (i.e., the probability that B will occur given that A has occurred), which can be corrected for chance (i.e., using the base rates of codes A and B). It is important to note that a null finding in the first example does not mean that children do not acquiesce following leading questions, simply that they do not acquiesce more often than chance, which would be expected if leading questions evoked acquiescence. A null finding would indicate that children do not acquiesce more or less often following leading questions than they do following neutral questions.

Virtually every suggestibility study has shown that, compared with neutral questions, misleading questions are more likely to provoke inaccurate acquiescence by a child (e.g., Ceci & Bruck, 1995). However, as mentioned previously, these data are based on a paradigm in which the interviewer’s questioning is not allowed to vary in response to the behaviors and characteristics of the child, and are based on analysis of adjacent pairs of turns in structured interviews without consideration of more distal chains. Thus, researchers do not know why in actual conversation streams interviewers choose to use leading questions at the precise locations they do or why they sometimes elect to repeat suggestions and at other times they do not. Perhaps, as Gilstrap and Papierno’s (in press) analyses suggest, interviewers are being influenced by the child. Interviewers may use leading questions with recalcitrant children who have earlier denied abuse, and therefore leading questions might be more likely to be followed by additional denial than by acquiescence. In unstructured interviews, do leading questions provoke acquiescence or affirm earlier denial?

By allowing the interviewers’ questioning to be responsive to children’s replies, researchers can examine child-to-adult effects, which would illuminate, for example, whether leading questions are more likely to be used following a child’s denial, or whether a child’s acquiescence is more likely to follow a certain pattern of prior exchanges. Do child behaviors affect adult behaviors?

With sequential analyses we can ask whether the antecedent of a behavior affects the consequent in ways that would be opaque within the traditional analytic framework. For example, does the consequence of a leading question (e.g., leading → ?) differ depending on the child’s previous response? It could be that leading questions are likely to produce acquiescence only when they follow earlier acquiescence (i.e., child acquiescence → interviewer leading → child acquiescence), and that leading questions that follow earlier denial are likely to produce additional denial (i.e., child denial → interviewer leading → child denial). If this proves to be the case, it implies that the very same type of interviewer utterance may be more or less deleterious, depending on the child behavior that preceded its use. A given type of utterance may be significantly more
detrimental in one place in the conversation than in another.

Whether sequential analysis proves useful in revealing characteristics of child suggestibility is an empirical question, the answer to which will depend on whether the data contain temporally contingent behaviors. That is, are children’s reporting behaviors contingent on interviewers’ questioning behaviors, and vice versa? In the following study we provide a direct test of this question in a contrived forensic situation, focusing on videotaped professional interviews about a staged event in which the interviewers were allowed to tailor their questioning to the child’s individual responses rather than adhere to scripted protocols that allow little variance.

**Method**

**Participants**

Participants were 41 preschool and primary school children ranging from 36 to 82 months (M = 56.4 months, SD = 11.6) with approximately half of the children below 5 years old and half above 5 years old. The sample consisted of 20 females and 21 males. Parents of children were contacted directly for permission for their children’s participation. Forty-one experienced interviewers (31 females, 10 males) who specialized in conducting child protection interviews participated in the study. Interviewers were recruited with the help of local agency heads. The mean age of the interviewers was 32.7 years (SD = 3.9 years). On average, 86.1% of the interviewers worked full time and interviewed 2.6 children per week (SD = 1.5). Interviewers averaged 11.8 years of interviewing experience (SD = 4.5 years) with an average of 2.9 years of experience interviewing children (SD = 1.9 years). All interviewers reported that they testified regularly in court. Both the sample of interviewers and the sample of children were British and were primarily of Northern European descent.

**Design and Procedure**

One month before the interviews were conducted, an event was staged in the children’s classrooms. A very similar event had been used in prior research and its components were found to be salient and exciting for children in this age range (Warren et al., 1999). The event entailed a visit by a magician while the children were engaged in a scheduled discussion during their regular circle time. The teacher stopped the discussion and asked children to pay attention to the visiting magician. The magician performed three script-consistent actions (e.g., typical card tricks) and three script-inconsistent actions (e.g., make the audience into an orchestra). Additionally, the magician wore clothing items that were both script-consistent (e.g., wearing black pants and red coat) and script-inconsistent (e.g., wearing an apron). The event was videotaped.

Beginning 1 week before the start of the interviews, researchers acclimated to the classroom. At the start of each interview session, interviewers were reminded of the name of the child and told that they had a maximum of 25 min to interview the children. Based on previous research, we expected the interviews to last 15 to 20 min (Gilstrap, Nunno, Thomas, & Toglia, 1999). The researcher informed the interviewers that they should interview the child “as they normally would, using the techniques that they would normally use” in their actual forensic work. Although none of the interviewers had seen the event, they were given basic information about the event in the same form they would receive in the field. They were told that a magician had visited the classroom, and they were given a list of six actions and told that they may or may not have occurred during the magician’s visit (two pieces of information were accurate, two were of mixed accuracy, and two were inaccurate) and to treat the information as they would an actual incident report—that is, with some skepticism. All interviews were both audiotaped and videotaped for transcription and coding.

**Transcription.** Staff trained by the police training facility to provide detailed verbal transcriptions as well as pertinent nonverbal behaviors, such as nods or gesticular descriptors, transcribed all videotaped interviews. The transcripts were then reviewed a second time by research assistants to check for accuracy. The resulting transcripts were printed and used in conjunction with the videotapes for the remainder of the coding process.

After transcription was completed, three separate sets of coding were conducted: interview stages, adult turn types, and child turn types. Coders were trained on transcripts from a separate data set until a minimum of 90% agreement was reached with the primary investigator. All interviews in the current sample were coded independently by two coders and any differences were resolved by discussion. An additional 10% of the interviews were coded separately by the principal investigator to calculate interrater reliability. Reliability was calculated at the most conservative level (i.e., on the most specific level of turn type-accuracy-redundancy string) using GSEQ software designed for observational coding (see...
Bakeman & Quera, 1995, for a description of this program). Even at this level of specificity, interrater reliability was satisfactory, with Cohen’s $\kappa = .74$ (for a discussion of why Cohen’s kappa is generally a more conservative measure of interrater reliability than percentage agreement, see Bakeman & Gottman, 1997).

Coding: Interview stages. First the interview was divided into stages: introduction, rapport building, topic of concern, free narrative, specific questions, and closure. The stages did not always occur in this order, and not all interviews had all stages. Only the free narrative and specific questioning stages, by definition, contained event-related discussion. These stages were divided into on-topic and off-topic portions. On-topic portions of free narrative and specific questioning were then coded for specific adult and child behaviors.

Coding: Adult turn types. Separate coders were used to code child and interviewer turn types for all sections of free narrative and specific questioning that were on topic. Turns were coded with two coding schemes, one for the interviewer and one for the child. The interviewer coding scheme was divided into leading and neutral turn types as follows.

Leading questions were defined as any question that introduced information into the conversation. Questions were considered leading when the interviewer stated a preference (e.g., I think the magician did X, do you?), repeated a question, introduced new information (e.g., Did he have a blue hat?), limited the range of possible answers by forcing choices (e.g., Was it red or blue?), asked the child to imagine or suppose or consider options (e.g., What might the magician have done?), or misphrased the child’s statement (e.g., child: “It was like Flubber”; interviewer: “Like Flubber?”). Because we defined leading according to strict logical decision rules, one additional question type was included among leading questions that many consider neutral. This type of question introduces the possibility that something occurred (e.g., What else did he do?) without suggesting what. If the person did not do anything else (or anything at all), such a question is inaccurately suggesting that he or she did. Also, note that because leading questions include information, they are more similar to recognition questions, and because neutral questions do not include information, they are more similar to recall questions. This is an important distinction but beyond the scope of the current work.

Neutral questions were defined as any interviewer statement that did not introduce information into the conversation. Neutral questions included those when the interviewer followed up on items that the child had previously mentioned (e.g., child: “He had a hat”; interviewer: “What color was the hat?” or, “Tell me everything about that”), made facilitative statements that contained no information related to the forensic theme (e.g., Was that fun? or, Really…), clarified the child’s language use (e.g., What is Flubber?), or correctly rephrased the child’s statement (e.g., child: “It was like Flubber”; interviewer: “Like Flubber?”).

In the adult coding scheme, because all leading questions contained some form of introduced information, they were also coded for the accuracy of the suggested information—accurate or inaccurate—and redundancy—whether the suggestion was being introduced for the first time (new) or had been introduced previously in some form (redundant). Originally, mixed-accuracy, and mixed-redundancy codes were also included but because of the extremely low base rate of these codes they were eliminated. Mixed-accuracy codes (containing some inaccurate and some accurate information) were considered inaccurate and mixed-redundancy codes (containing some new suggested information and some already suggested information) were considered new.

Coding: Child turn types. The coding scheme for the children’s responses, verbal or nonverbal, was coded globally into acquiescence (some level of positive response to the interviewer’s question) and denial (some level of negative response to the interviewer’s question). Because child coders could see the adult codes, which were entered first, two-letter codes were used so that the meaning of the code would not be understood. Thus, child coders were unaware of the codes the adult coders had given and vice versa.

Acquiescence was defined as giving details that addressed the interviewer’s question and expressing agreement. Denial was defined as expressing disagreement, giving details that departed from the topic of the interviewer’s question, saying “I don’t know,” giving no response, and giving no information.

Child codes were also coded for accuracy (whether the response was accurate or inaccurate with regard to the event) and redundancy (whether the response gave new or redundant information). To give children the benefit of the doubt and to avoid unfairly inflating estimates of inaccuracy, all details not related to the event, I don’t know statements, lack of response, and other statements were automatically coded as accurate even though surely some were not.

Results

Analytic Plan

We begin our analyses by presenting base rate information for all behaviors. We then continue with our
three major sequential questions: Are leading questions followed by acquiescence in field interviews? Are child behaviors predictive of interviewers’ behaviors? Does the accuracy and redundancy of the information in the question affect children’s responses?

After answering these questions, two post hoc questions arose: (a) Are children and adults remaining consistent such that it is possible to predict from a behavior to their next behavior, ignoring the other conversational partner’s intervening turn? Is there any evidence of hierarchy such that children or adults are more consistent or more variable? (b) Can the effect of new misleading questions be explained in the same manner as leading questions generally; that is, are new misleading questions provoked by child behaviors and thus the consequences of these questions are in actuality the result of children remaining consistent?

**Base rates.** We first present nonsequential base rate data for the variables on which we performed sequential analyses. Several different classes of scores were computed from the observational data. Frequency data represent mean number of behaviors per interview for the behavior being discussed (except where percentages are noted).

**Simple two-code sequences.** Two-code sequences are the heart of sequential analysis; they allow us to ask what the transitional probability is between two potentially temporally linked behaviors. We began our sequential analyses with the most general level of two-code sequences—interviewer-to-child influences (e.g., an interviewer’s leading questions → a child’s denial) and child-to-interviewer influences (e.g., a child’s denial → an interviewer’s subsequent leading questions). We start with global codes (lead vs. neutral and acquiescence vs. denial) and then explore changes in these global findings when subcodes are added (accuracy and redundancy).

For any sequential data we computed transitional probabilities that a specific event (the target event) will follow another event (the given event), correcting for the base rate of the given event (see Bakeman & Gottman, 1997; Bakeman & Quera, 1995, for detailed discussion of these methods). Instead of transitional probabilities, we present Yule’s Q, which accounts for base rates of the target event and is less affected by the number of tallies, as it has been found to be the best measure of effect size for these types of contingency tables (Bakeman, McArthur, & Quera, 1996). Yule’s Q is an effect size measure and can be interpreted in a similar manner as correlations, ranging from −1 to 1, with effect strength increasing as you approach 111. Yule’s Qs were computed for individual dyads. Means for all interviews are presented as well as a sign test indicating whether the majority of dyads are in the same direction as the mean.

It is important to note that sequential analysis deals with changes in the base frequency of the target behavior that depend on the given behavior. Thus, if acquiescence (a potential target behavior) had a base rate of 40% across interviews and occurred 40% of the time following leading questions (a potential given behavior), this means there is not a temporal or conditional relationship between leading questions and acquiescence. This does not mean that acquiescence does not follow leading questions at a level that we should be concerned about for applied reasons. In this example, acquiescence follows leading questions 40% of the time, which is certainly a forensic concern. However, if researchers are interested in causal mechanisms, it fails to support an evocative or causal relationship between leading questions and acquiescence.

**Lagged analyses (post hoc).** As will be seen, strong relationships within the two-code sequences (adult-to-child and child-to-adult) led us to ask whether, based on these initial findings, we could predict directly from adult-to-adult behavior as well as from child-to-child behavior, effectively skipping the intervening behavior of the other conversational partner. Using lagged analyses we asked whether there was a conditional relationship between an adult’s behavior and the adult’s next behavior (and the same for child behaviors). These analyses allowed us to ask questions of the type: Does a behavior that occurred two codes previously (e.g., a child’s acquiescence) exert influence on the code of interest (e.g., a child’s current acquiescence) regardless of the intervening code (e.g., an interviewer’s leading question)? Such analyses may allow us to unearth causal paths that are masked when focusing on local two-code chains of interviewer and child behaviors.

**Base Rates**

Means and standard deviations for all codes are presented in Table 1 (for a more detailed presentation of interviewers’ behaviors in the same sample, see Gilstrap, 2004). In this sample, we find that 42% of adult questions were leading and 58% were neutral. Of the leading questions, 72% suggested accurate information and 28% suggested inaccurate information, and 67% suggested new information to the child and 33% suggested old (already suggested) information. The most common type of leading question was new, accurate suggestion (46%), followed in frequency by redundant, accurate suggestion (26%); new, inaccurate suggestion (21%); and
redundant, inaccurate suggestion (7%). For children’s behaviors, we found that children acquiesced 57% of the time and denied 43% of the time. Child acquiescence and denial were more often accurate (73% and 89%, respectively) than inaccurate (27% and 11%) and more often new (77% and 90%) than redundant (23% and 10%). The most common type of child response was an accurate, new response both for acquiescence (58%) and denial (82%).

Child age was not correlated with total leading questions asked, total neutral questions asked, total denials given, total acquiescences given, or total turns (see Table 2 for age and gender analyses). A series of one-way analyses of variance (ANOVA) predicting total leading questions, neutral questions, denials, and acquiescences from gender were performed. A gender difference was found for acquiescences given, $F(1,40) = 4.11$, $p < .05$, with boys ($M = 33.0$, $SE = 5.0$) giving more acquiescences than girls ($M = 27.4$, $SE = 4.7$).

### Simple Two-Code Chains

**Global-level analyses.** We first examined the probability that an interviewer’s leading statement would be followed by a child’s acquiescence, as would be expected based on the literature reviewed in the Introduction (see Table 3). Children were less likely to acquiesce following a leading question than they were to deny the leading question (mean $Q = -.20$), and conversely they were more likely to acquiesce following a neutral question and less likely to deny.

Regarding the second major question, we found that children were being asked leading questions when they were denying. Overall, acquiescence by children was unlikely to be followed by leading questions (mean $Q = -.42$) and more likely to be followed by neutral questions. Conversely, denial was likely to be followed by leading questions and unlikely to be followed by neutral questions.

**Micro-level analyses.** When the accuracy and redundancy of the questions and answers were considered (see Table 3, beginning in third row), the number of tallies per cell was a concern. We eliminated any interviews where either the target or given code occurred less than 5 times and confined gender and age analyses to behavior chains where more than 20 interviews remained for analyses. Overall, in this reduced sample the sublevel analyses reflected results that were either in a similar direction to the global findings or were nonsignificantly different from zero. Sign tests indicated that in our sample new, accurate leading questions were less likely than chance to be followed by redundant, accurate acquiescences ($Q = -.49$); new, misleading questions were unlikely to be followed by new accurate acquiescences ($Q = -.40$), and redundant, accurate leading questions were unlikely to be followed by new, inaccurate acquiescence ($Q = -.75$). One notable exception to the global pattern was that new, misleading questions were more likely to be followed by new, inaccurate acquiescences than would be expected by chance ($Q = .27$). No significant trends were found for child subcode analyses following neutral questions, although mean Qs for acquiescence codes were positive and the mean Q for the remaining denial code was zero.

**Gender and age.** In the Lag 1 analyses, two gender effects emerged. Gender significantly predicted children’s likelihood to acquiesce to leading questions ($F = 4.10$, $p < .05$), with boys being significantly more likely to acquiesce ($M = -.07$, $SE = .09$) than girls ($M = -.33$, $SE = .10$). Gender also significantly predicted children’s likelihood of providing new, accurate acquiescences when asked new, accurately leading questions ($F = 7.60$, $p < .05$). Again, boys were more likely to acquiesce ($M = .31$, $SE = .08$)

---

**Table 1**

**Mean (and Standard Deviations) Speech Turns for Adult and Child Behaviors**

<table>
<thead>
<tr>
<th>Code</th>
<th>Adult Acquiescence</th>
<th>Adult Denial</th>
<th>Child Acquiescence</th>
<th>Child Denial</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (and Standard Dev.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leading New</td>
<td>9.8 (1.0)</td>
<td>5.5 (7.2)</td>
<td>15.3 (1.4)</td>
<td>8.6 (1.7)</td>
</tr>
<tr>
<td>Leading Redundant</td>
<td>5.5 (7.2)</td>
<td>1.5 (4.4)</td>
<td>6.1 (1.0)</td>
<td>2.2 (0.5)</td>
</tr>
<tr>
<td>Leading Total</td>
<td>15.3 (1.4)</td>
<td>6.1 (1.0)</td>
<td>21.3 (2.1)</td>
<td>20.8 (1.9)</td>
</tr>
<tr>
<td>Neutral New</td>
<td>29.3 (3.0)</td>
<td>18.6 (1.7)</td>
<td>20.1 (1.9)</td>
<td>17.5 (1.9)</td>
</tr>
<tr>
<td>Neutral Redundant</td>
<td>5.5 (7.2)</td>
<td>0.8 (2.0)</td>
<td>2.5 (0.5)</td>
<td>0.8 (0.3)</td>
</tr>
<tr>
<td>Neutral Total</td>
<td>29.3 (3.0)</td>
<td>18.6 (1.7)</td>
<td>32.8 (2.3)</td>
<td>20.3 (2.6)</td>
</tr>
<tr>
<td>Denial New</td>
<td>17.5 (1.9)</td>
<td>20.1 (1.9)</td>
<td>4.7 (4.9)</td>
<td>22.2 (2.5)</td>
</tr>
<tr>
<td>Denial Redundant</td>
<td>5.5 (7.2)</td>
<td>2.2 (0.5)</td>
<td>4.7 (4.9)</td>
<td>8.1 (1.4)</td>
</tr>
<tr>
<td>Denial Total</td>
<td>22.2 (2.5)</td>
<td>8.1 (1.4)</td>
<td>30.2 (3.4)</td>
<td>30.1 (3.4)</td>
</tr>
</tbody>
</table>

**Table 2**

**Correlation Differences M (SE) for Gender**

<table>
<thead>
<tr>
<th>Code</th>
<th>Gender Correlation Differences M (SE) for Age</th>
<th>Females</th>
<th>Males</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leading questions</td>
<td>-.05</td>
<td>2.39</td>
<td>18.1 (2.7)</td>
</tr>
<tr>
<td>Neutral questions</td>
<td>.13</td>
<td>1.22</td>
<td>25.9 (3.8)</td>
</tr>
<tr>
<td>Child denial</td>
<td>-.12</td>
<td>0.66</td>
<td>18.4 (2.7)</td>
</tr>
<tr>
<td>Child acquiescence</td>
<td>.14</td>
<td>4.11</td>
<td>27.4 (4.7)</td>
</tr>
<tr>
<td>Total turns</td>
<td>.06</td>
<td>2.27</td>
<td>89.7 (11.0)</td>
</tr>
</tbody>
</table>

$p < .05$.
Table 3
Lag 1 Analyses

<table>
<thead>
<tr>
<th>Given</th>
<th>Target</th>
<th>n</th>
<th>Mean transitional probability</th>
<th>Mean Q</th>
<th>n in same direction</th>
<th>r with age</th>
<th>Females’ mean Q (SE)</th>
<th>Males’ mean Q (SE)</th>
<th>Gender F test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leading</td>
<td>Acquiescence</td>
<td>41</td>
<td>.50</td>
<td>−.20</td>
<td>29/41*</td>
<td>.02</td>
<td>−.33 (.10)</td>
<td>−.07 (.09)</td>
<td>4.10*</td>
</tr>
<tr>
<td>Acquiescence</td>
<td>Leading</td>
<td>41</td>
<td>.57</td>
<td>−.42</td>
<td>38/41*</td>
<td>−.02</td>
<td>−.48 (.10)</td>
<td>−.36 (.10)</td>
<td>0.74</td>
</tr>
<tr>
<td>Neutral</td>
<td>Acquiescence – accurate – new</td>
<td>34</td>
<td>.38</td>
<td>.02</td>
<td>16/33</td>
<td>.07</td>
<td>−.09 (.08)</td>
<td>−.04 (.08)</td>
<td>1.41</td>
</tr>
<tr>
<td>Neutral</td>
<td>Acquiescence – inaccurate – new</td>
<td>17</td>
<td>.19</td>
<td>.06</td>
<td>10/17</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neutral</td>
<td>Acquiescence – accurate – redundant</td>
<td>12</td>
<td>.20</td>
<td>.21</td>
<td>7/9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neutral</td>
<td>Denial – accurate – new</td>
<td>39</td>
<td>.40</td>
<td>.00</td>
<td>22/38</td>
<td>−.15</td>
<td>−.02 (.10)</td>
<td>.02 (.09)</td>
<td>0.09</td>
</tr>
<tr>
<td>Leading accurate new</td>
<td>Acquiescence – accurate – new</td>
<td>25</td>
<td>.43</td>
<td>.15</td>
<td>16/24</td>
<td>−.18</td>
<td>−.02 (.09)</td>
<td>.31 (.08)</td>
<td>*7.60</td>
</tr>
<tr>
<td>Leading accurate new</td>
<td>Acquiescence – inaccurate – new</td>
<td>17</td>
<td>.17</td>
<td>−.09</td>
<td>10/17</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leading accurate new</td>
<td>Acquiescence – accurate – redundant</td>
<td>12</td>
<td>.08</td>
<td>−.49</td>
<td>11/12*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leading inaccurate new</td>
<td>Acquiescence – accurate – new</td>
<td>12</td>
<td>.21</td>
<td>−.40</td>
<td>10/12*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leading inaccurate new</td>
<td>Acquiescence – inaccurate – new</td>
<td>10</td>
<td>.33</td>
<td>.27</td>
<td>9/10*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leading inaccurate new</td>
<td>Acquiescence – accurate – redundant</td>
<td>4</td>
<td>.05</td>
<td>−.51</td>
<td>3/4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leading inaccurate new</td>
<td>Denial – accurate – new</td>
<td>14</td>
<td>.41</td>
<td>.00</td>
<td>5/14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leading accurate redundant</td>
<td>Acquiescence – accurate – new</td>
<td>17</td>
<td>.30</td>
<td>−.14</td>
<td>11/17</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leading accurate redundant</td>
<td>Acquiescence – inaccurate – new</td>
<td>11</td>
<td>.05</td>
<td>−.75</td>
<td>10/11*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leading accurate redundant</td>
<td>Acquiescence – accurate – redundant</td>
<td>7</td>
<td>.09</td>
<td>−.49</td>
<td>4/7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leading accurate redundant</td>
<td>Denial – accurate – new</td>
<td>22</td>
<td>.43</td>
<td>.01</td>
<td>11/20</td>
<td>−.31</td>
<td>.01 (.15)</td>
<td>.02 (.13)</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Note. Analyses only included interviews in which the base rate for both the target and the given exceeded 5. If n ≥ 20, correlations were reported, and if n for males and females were both ≥ 10, gender tests were conducted. Gender F tests covaried age. All sequential analyses in this table were performed at Lag 1. Percentage in same direction is out of total interviews minus ties (zeros). Significance values on n with positive Q reflect a one-way sign test. The sign test was conducted to eliminate the hypothesis that a few extreme values (interviews) could be influencing the mean. All tests of significance are one-tailed in the direction of the observed effect. The resulting p values represent the probability that the true value of the effect is of sign opposite to the observed value.

* p < .05.
The findings from the present study demonstrate the usefulness of sequential analyses in the study of the consistency effect. The average effect size for children was larger than that of adults, and individual dyads conformed to the child effect direction reliably (p < .001) whereas individual dyads did not conform to the adult effect direction reliably (p = .06).

We conducted a second set of Lag 2 analyses on the one exception we found to our global analyses when considering accuracy and redundancy. We had previously found that new misleading questions asked by adults were likely to be followed by inaccurate new acquiescences by children. Next, we asked whether new misleading questions were preceded by inaccurate new acquiescences. Taken in combination with our Lag 2 findings regarding global child consistency, this would lead us to wonder if the new inaccurate acquiescences were a consequence of prior new inaccurate acquiescences. However, in this case, the child behavior did not predict the adult behavior; that is, inaccurate new acquiescences by children were not reliably followed by new misleading questions when compared with chance (Q = .05, p = .75). Because of the importance of a potential relationship between acquiescence and leading questions, with reference to the suggestibility studies that have examined this relationship, and with reference to the real-world importance of child acquiescence to misleading suggested information, we performed the Lag 2 analyses on child inaccurate, new acquiescence and on interviewer inaccurate, new suggestion. Sign tests indicated that neither the new, misleading question Lag 2 analyses nor the new, inaccurate acquiescence Lag 2 analyses were significantly different from zero (Qs = .11 and .02, respectively).

**Discussion**

The findings from the present study demonstrate the usefulness of sequential analyses in the study of...
child suggestibility and interviewer effectiveness more generally. A number of results emerged that would not have been evident using traditional analyses, as we discuss now.

There were three results that were initially counterintuitive and carried unanticipated forensic as well as theoretical implications. First, analyses of unstructured interviews in which field interviewers are able to tailor their line of questioning in response to the child’s response style and personality appear to yield results that are, at least at the global level, at variance from those produced by studies of scripted, highly structured interviews. Unlike the results of these highly structured studies, we found that overall leading questions were likely to be followed by denial, not acquiescence. Such a result calls into question the customary finding that interviewers’ suggestive questions are more likely to be followed by increased inaccurate acquiescence by children. Having stated this, however, we also found that when specific code chains were examined, new misleading questions predicted new inaccurate acquiescence. This could be the difference between our field sample and structured interviews. In structured interviews, researchers ask exclusively new misleading questions. We return to these contradictory findings in our discussion of the primacy of delayed child-to-child influences (i.e., child consistency).

Second, we asked what types of child behaviors were prompting interviewers to pose suggestive questions, a heretofore unstudied relationship. We found that child denial was more likely to be followed by suggestive questioning than was child acquiescence. The idea that interviewers use leading questions with unresponsive children is a much hypothesized and warned about link that had not been previously tested. It has been suggested that interviewers might use leading questions inappropriately to obtain “coerced” confessions from recalcitrant children. Indeed, one might ask why we would expect an interviewer to use a leading question with a child who is already acquiescing. Although our interviews did not contain the same level of coercive behavior found in some highly publicized criminal investigations, we did find that interviewers followed denial with leading questions. This led us to hypothesize that adults were using leading questions only with recalcitrant children, and thus the relationship between leading questions and subsequent child denial might actually be driven by the child denial that preceded.

Thus, the third counterintuitive finding was that for both leading and neutral questions, the child’s behavior that preceded an interviewer’s question was predictive of the child’s subsequent response to that interviewer’s question regardless of the type of question asked by the interviewer. Using lagged sequential analytic techniques, we found that we could predict from initial child denial to subsequent child denial, effectively ignoring the interviewer behavior. This was a startling result because it appears to suggest that child behavior is not affected by interviewer behavior at all. In accord with this assumption, we found that we could not reliably predict from interviewer behavior to interviewer behavior, effectively skipping the intervening child behavior. An implication of this finding is that leading questions seem to be less important in evoking acquiescence from children than their own prior behavior. Perhaps interviewers employed leading questions in a manner that took into account the child’s personality and response style, and thus used leading questions to dislodge stubborn denials rather than to pry loose acquiescence from cooperative children. Recall that the use of leading questions was not generally related to acquiescence in the current interviews. (Indeed, they were related to decreased acquiescence, as evidenced by the negative coefficient for Yule’s Q.) It could be argued that a child’s denial is simply a response to the consistent use of acquiescence, as evidenced by the negative coefficient for Yule’s Q. Other support for the primacy of the influence of the child comes from Gilstrap and Papierno (in press). Although these lagged analyses do not preclude the presence of a significant relationship between interviewer behaviors and child behaviors, they suggest that the child’s behavior was more predictive of his or her subsequent narrative recall performance than the interviewer’s behavior. Theoretically, this is perhaps the biggest surprise in these results, and the one we expect will spawn the most activity to unravel.

This allows us to return to a possible explanation of why, globally, we find that leading questions are followed by denial but in some specific instances we find that leading questions are followed by acquiescence. Because denial is likely to be followed by leading questions, generally when we are studying the effects of a leading question we are actually likely to be studying a denial → leading chain. Thus, with the primacy of the delayed child effects, we would expect that denial would be followed by denial, and thus, in most instances leading questions would be followed by denial because of the delayed child effect not because of the interviewer effect. The fact that despite the child consistency, new misleading questions were related to new inaccurate
acquiescence supports findings from the literature on leading questions and acquiescence using structured interviews. We asked whether new misleading questions were preceded by new inaccurate acquiescence, and thus when we were studying new leading questions we were actually studying the new inaccurate acquiescence → new misleading question chain. However, new inaccurate acquiescences were not more likely to precede new misleading questions. Therefore, this explanation does not fit and an explanation for why new misleading questions are unique awaits further investigation.

In several instances our findings were moderated by gender differences. Girls were less likely to acquiesce overall, were less likely to acquiesce after leading questions, and were less likely to acquiesce inaccurately after new misleading questions. Gender findings in suggestibility are not ubiquitous. Of the few articles we found that included gender differences, they were relatively small in magnitude and showed that girls were more likely to acquiesce than boys (Calicchia, Santostefano, & Calicchia, 2004; Ghetti, Goodman, Eisen, Qin, & Davis, 2002). The fact that girls were more likely to follow our overall pattern of findings than boys warrants further study. It is important to note, however, that the majority of actual forensic interviews are conducted with girls.

As alluring as the claim might be that leading questions are not the culprit in unstructured interviews that they have been shown to be in scripted interviews, we cannot conclude that interviewers’ use of leading questions is simply a response to a child’s denial. First, because either the behaviors of the interviewers nor the children were randomly assigned in this study, we are constrained in pressing this point. Additionally, the design may have mitigated the effects of suggestions by virtue of informing the interviewers that they would be videotaped and their performance would be analyzed. This could have resulted in some restraints on the nature of the suggestive questions they used. We do not have evidence of this, but such a possibility must be acknowledged, as we discuss later. Finally, the gravity of forensic events may be so unlike the magician event used in the present analysis that they may render children’s willingness to acquiesce and deny differentially susceptible to interviewer pressure. Even if true, however, the present results provide strong evidence for the value of using sequential analytic methods with actual frontline interview transcripts.

Caveat Lector

In view of the counterintuitive nature of some of the preceding results and their forensic implications, we wondered whether we might have missed something. At first glance, the results of the present investigation seem to confirm some prior research findings while simultaneously casting doubt on others. Specifically, the present findings demonstrate that overall leading questions do produce denial, but when the analyses are confined to inaccurate new leads, they produce acquiescence, as would be expected from the bulk of the laboratory research. Hence, we show that inaccurate new leads, the kind used in structured interviews, do produce acquiescence on the child’s part in unstructured interviews. Having said this, we also found that, overall, the child’s own prior response was a primary influence in producing his or her subsequent response, without consideration of the intervening questioning by the interviewer. How can both of these findings be correct? In particular, we considered three explanations.

First, perhaps a causal link between suggestions and inaccurate acquiescence did in fact exist early in the interview but was masked by a lack of relation downstream in the conversation. In other words, perhaps the first few leading questions used early in the interview provoked a jump-start of acquiescence but later leading questions, perhaps used as follow-up probes, provoked denial. These first few leading questions could have strong effects simply by starting the initial flow of the conversation in a given direction. To examine this we explored the relationship between the first leading question and acquiescence and denial. The first suggestive question occurred on average between the third and fourth interviewer on-topic turn. This analysis revealed that the first suggestive question led to, if anything, more denying, although our sample for this analysis was significantly diminished (i.e., the same number of interviews but only one question, the first leading question, per interview). The first leading question was followed by denial 65% of the time and acquiescence the remaining 35%. In the full-interview data, leading questions were followed by denial 51.4% of the time and acquiescence 48.6% of the time. Notwithstanding this limited sample, there is no support for the view that we missed a significant link between interviewer’s suggestions and children’s acquiescences earlier in the interview.

Second, we considered the possibility that our definition of leading questions was overinclusive, inflated by the inclusion of weakly or nonsuggestive leading questions. Because of the exploratory nature of the second and third questions, we present only descriptive data on the effect sizes for each condition pooling across dyads without values that could be inflated by the increased sample size (i.e., z scores or


...p values). If true, this could have resulted in an underestimation of the direct causal link between questions that are strongly suggestive and children's acquiescence to them. Specifically, in the analyses reported, leading questions included many examples that might not be regarded as strongly suggestive, such as when an interviewer merely asked, "What else did he do?" in a situation where the magician had done nothing else, or when an interviewer introduced new information but in a manner that did not suggest a desired answer (e.g., Did he have a hat on?). We wanted to know children would deny when the interviewer's questions were strongly suggestive, such as when the interviewer states a specific preference (e.g., I think the magician did X, don't you?), or repeats a question, or misphrases the child's statement. According to Ceci and Friedman (2000), such strongly suggestive utterances are of the most concern, not the mild leads mentioned earlier. To explore this possibility, we conducted one final analysis: We asked what was likely to follow and precede such strongly leading questions. We restricted the analysis to two-code behavior chains in which the interviewer employed strongly leading questions to determine whether such strong suggestions predicted child acquiescence. We discovered that these strongly suggestive utterances by the interviewer were equally (if not more) likely to provoke denial and to be provoked by denial. Strong leading questions were likely to be followed by denial and not by acquiescence whether the strong leading questions were inaccurate (Q = .42) or accurate (Q = .42) and perhaps even more strongly related to increased denial than weak leading questions (Q = -.15). All types of leading questions were more likely to follow child denial (strong leading inaccurate, Q = .41; strong leading accurate, Q = .60; and weak leading, Q = .25). Of course, this raises the possibility that children perceived the interviews as game-like, allowing them to disagree with a good-natured adult about pointedly phrased questions. This might account for the rather anomalous finding that strongly inaccurate suggestions by the interviewer were actually more likely to prompt denial than were weakly worded suggestions. Only further research can determine whether this is an issue.

Third, we considered differences between the codes that make up the overall leading category versus the inaccurate new leading category. In this data set, repeated questions make up 13% of overall leading questions (Gilstrap, 2004) but are not included in estimates of inaccurate new leading because they are redundant by definition. It could be that the increased denial following leading questions is being driven by the repeated questions code and that is why inaccurate, new leads, which do not include repeated questions, differ by predicting acquiescence. The answer is that repeated questions are the strongest predictor of denial but the other three most frequently occurring codes that are all included in inaccurate, new leads also predict denial. Preference (Q = -.24), question repetition (Q = -.54), introduced information (Q = -.11), and introduced possibility (Q = -.18) all led to decreased acquiescence. No significant relationships existed for misphrase and imagination requests, which occurred infrequently.

**Forensic Implications**

As already noted, our interviewers were aware that they were being taped, and they may have felt less pressure to obtain a disclosure than would be the case in investigative and lab studies. Perhaps these interviews were not "effective" in the sense that a certain degree of suggestibility was neither sought nor needed. If a laboratory researcher discovered that leading questions did not result in children's acquiescence, this might result in redesigning the study to include stronger suggestion to make certain there was sufficient acquiescence to test their hypothesis. For example, if one is interested in whether 4-year-olds are more suggestible than 6-year-olds, and pilot testing revealed that there was no link between leading questions and acquiescence, before the researcher would conclude there were no age differences they might first question whether the manipulation was too weak to detect age differences that would be present with stronger leads, longer latencies, the inclusion of stereotypes, and so on. There is a validity to such redesigns: Imagine if a suggestibility researcher asked all children questions suggesting that their names were incorrect, that their real names were other than they knew to be true. None of the children would succumb to such suggestions, of course. But no researcher would conclude from this result that children are not suggestible under other circumstances. By analogy, perhaps the present interviews, although effective and useful for our purposes, were not as suggestive as those used in studies that find that interviewer's leading questions predict children's acquiescence.

In sum, our results should not be construed to refute traditional notions about the direct causal influence of suggestive utterances by interviewers. Instead, although the findings converge on the view that it is the child's own behavior that is primary, this may not be the case in more suggestive contexts such as those reported in the literature where a leading
question elevates the likelihood of incorrect acquiescence or in criminal investigations where interviewers may have incentives to obtain disclosures.

However, such a view leaves unanswered two questions that cannot be addressed with the present data. First, it leaves open the issue of what determines the child’s behavior, if not the interviewer’s questioning. Given that numerous studies indicate that suggestive questions can be deleterious, if such suggestive questions are not the culprit, why do children respond differentially to variable degrees of suggestion? Second, several studies, including those conducted in unstructured field settings, have demonstrated that when interviewers employ suggestions, children are significantly more likely to provide disclosures than when interviewers do not employ these suggestive techniques (Lamb et al., 1996). We recategorized our data according to informativeness (i.e., providing information even if through denial) instead of acquiescence (i.e., providing information in line with the suggestions) as similarly as we could to Lamb, Sternberg, and Esplin (2000) given different initial coding schemes. As they found with their data, we also found that leading questions produced more new details. Suggestive questions (i.e., preference statements, repeated questions, and introduced information) were more likely to be followed by new details (i.e., all event-related details, whether or not question related, and denying and assenting) than neutral questions (Q = .22).

Such findings militate against the view that the interviewer’s behavior has no primary role in producing errors or has a role but one that is honed to the child’s own response style and, therefore, is not likely to cause the sorts of problems that inhere when interviewers employ the same questions to all children. We do not conclude that leading questions are without causal impact. Instead, we argue that in unstructured interviews, leading questions tend to be used with children who are already denying, and because of what appear to be strong consistencies in child behavior, children who deny are likely to continue to deny. Rather than arguing that leading questions are benign, we argue that leading questions must be considered in the context in which they are being asked, which includes the effects of the child on the adult. It could be that adult questions merely mediate child behaviors, or that children’s denials set in motion a chain of behaviors including the interviewers’ leading questions and that those questions still have some causal impact. Finally, bear in mind that the analyses undertaken were predicated on changes in the conjoint probabilities of the occurrence of B (e.g., acquiescence) given that A (e.g., leading questions) already occurred, corrected for base rates. Thus, a null finding does not mean that acquiescence does not follow leading questions but, rather, that children do not acquiesce more or less often following leading questions than they do to, say, neutral questions. It is possible for a suggestive question to be damaging when followed by acquiescence even if other child behaviors are even more predictive of acquiescence.

Our sequential analyses have provided a wealth of information regarding the interaction between interviewer and child behaviors in unstructured interviews but have also resulted in many unanswered questions. It seems likely that the resolution of these riddles will require additional experimental designs that permit alternative explanations to be ruled out—that is, with random assignment of questions across children, variation in interviewer incentives to obtain acquiescence, increased gravity of the event in the eyes of the children, and so forth. For now, we have to satisfy ourselves with showing that sequential analysis of temporal data in unstructured interviews is capable of providing insights and generating hypotheses that would not be generated with traditional nonequivalent approaches. And the use of such an analytic framework promises to contribute to both theory (e.g., directional and bidirectional influences) and practice.

References


