

Children's Reasoning About the Causal Significance of the Temporal Order of Events

Teresa McCormack
Queen's University Belfast

Christoph Hoerl
University of Warwick

Four experiments examined children's ability to reason about the causal significance of the order in which 2 events occurred (the pressing of buttons on a mechanically operated box). In Study 1, 4-year-olds were unable to make the relevant inferences, whereas 5-year-olds were successful on one version of the task. In Study 2, 3-year-olds were successful on a simplified version of the task in which they were able to observe the events although not their consequences. Study 3 found that older children had difficulties with the original task even when provided with cues to attend to order information. However, 5-year-olds performed successfully in Study 4, in which the causally relevant event was made more salient.

Much of the recent research on the development of causal reasoning has emphasized young children's competence with various aspects of causal reasoning (e.g., Corrigan & Denton, 1996; Gopnik, Sobel, Schulz, & Glymour, 2001; Schlottmann, Allen, Linderoth, & Hesketh, 2002). It has been well established that even in infancy children show some type of sensitivity to the causal relationships between events (Leslie, 1982; Oakes, 1994; Oakes & Cohen, 1990) and that children in the preschool years show an appreciation of the causal powers of familiar objects (Bullock, Gelman, & Baillargeon, 1982; Gelman, Bullock, & Meck, 1980) and seem to infer the causal powers of novel objects in principled ways (Shultz, 1982; Shultz & Kestenbaum, 1985). Indeed, the idea that relatively sophisticated causal reasoning abilities are intact early in development has played an important role in the influential "theory-theory" approach to conceptual development.

This emphasis on the competence of preschoolers' causal reasoning can be contrasted with a claim by Povinelli, Landry, Theall, Clark, and Castille (1999) that 3-year-olds have difficulty with a fundamental causal ability, a causal ability that Povinelli et al. have linked closely with the development of a concept of time. Specifically, they have argued that children of this age may not grasp that events that occurred in the more recent past may be more relevant to determining the current state of the world than events that took place at an earlier point in time. To use a common

example, imagine that you have lost your car keys and are trying to find them. You may mentally retrace the day's events in an attempt to remember what you were doing when you last had the car keys, because you know that this most recent event is the most causally relevant to the current location of the keys. Povinelli et al.'s claim is that 3-year-olds do not have this kind of understanding of the causal relevance of the relative recency of events.

They examined this ability in 3- and 5-year-olds (Povinelli et al., 1999, Study 5) using a delayed video feedback technique. Along with an experimenter, children took part in two different games, one after another, in a room in which there were two different colored boxes attached to the wall behind the location where the child was seated. During Game 1, and unbeknownst to the child, a second experimenter placed a puppet in, for example, the blue box; during Game 2, the puppet was moved, again unbeknownst to the child, to, for example, the red box. After the children had played both games, they were shown videotapes of themselves playing the games and were able to see that the second experimenter had placed the puppet first in one box and then in the other box located behind them. After viewing the two videotapes, children were asked to identify the current location of the puppet.

The crucial manipulation in Povinelli et al.'s (1999) study was that children were not necessarily shown the videotapes of the games in the order in which the games had actually taken place. Half of the children were shown the tape of Game 1 before that of Game 2, but the other half saw the tape of Game 2 before they were shown that of Game 1. Povinelli et al. wanted to examine whether children grasped that it was the more recent of the two events (i.e., the placing of the puppet in the red box during Game 2) that was crucial in working out where the toy was currently located. They found that 3-year-olds were unable accurately to identify the current location of the toy and that even a 5-year-old group was not significantly above chance at choosing the correct box. In a subsequent study (Study 6), the order in which the games were played was emphasized while the children were shown the videotapes. In that study, 5-year-olds were successful but 3-year-olds still performed at chance.

Teresa McCormack, Department of Psychology, Queen's University Belfast, Belfast, Northern Ireland; Christoph Hoerl, Department of Philosophy, University of Warwick, Coventry, England.

This research was supported by Grant R000 23 9002 from the United Kingdom (U.K.) Economic and Social Research Council and Grant 88/515050 from the U.K. Biotechnology and Biological Sciences Research Council, both awarded to Teresa McCormack. We are grateful to Naumana Amjad, Elmaire Linehan, and Victoria Bell for assistance with data collection.

Correspondence concerning this article should be addressed to Teresa McCormack, Department of Psychology, Queen's University Belfast, David Keir Building, Malone Road, Belfast, Northern Ireland BT9 5BP, United Kingdom. E-mail: T.McCormack@qub.ac.uk

Povinelli et al. (1999) described young children's difficulties with this task in terms of a failure to understand that the recent past is causally bound to the present. Perhaps a more general way to describe the nature of the difficulty is in terms of the idea that children fail to grasp the causal significance of the temporal order in which past events took place in working out the current state of the world. In what follows, we distinguish between two quite different ways in which children may be sensitive to temporal-causal relationships. We label the first type of sensitivity *temporal updating*. This is an important but relatively primitive ability that is insufficient to perform well on the Povinelli et al. task. The second type of sensitivity is more sophisticated, and we label it the ability to make *temporal-causal inferences*.

Temporal updating simply involves altering one's model of the world as a series of causally related events unfolds (McCormack, 2001). Even very young children are clearly capable of changing their model of the world as they observe or infer it being modified (demonstrated, for example, in successful performance on the Piagetian A-not-B and invisible displacement tasks). Further, preschool children can find out about a series of events subsequent to their occurrence (e.g., by being told about them) and update their model of the world accordingly. For example, imagine two events, E1 and E2, that are causally related such that the subsequent state of the world is affected by the order in which E1 and E2 occur. That is, one outcome obtains if E1 comes first followed by E2, and a different outcome obtains if E2 is followed by E1. (In the case of Povinelli et al.'s task, E1 could be the placing of the puppet in the blue box and E2 the placing of the puppet in the red box.) Providing the child finds out about the events in the order in which they occur or have occurred, they could update their model of the world first in response to E1, then in response to E2, and thus would come to an accurate conclusion about the overall outcome. That is, they need only update their model of the world sequentially as they find out about each event.

Although simple updating will be sufficient in some circumstances, it will be insufficient, for example, in circumstances when the order in which children find out about events in the series differs from the order in which the events actually occurred. This is the case in Povinelli et al.'s (1999) study, in which the use of videotapes allowed the experimenters to manipulate the order in which children found out about the two hiding events, such that it was not necessarily identical to the order in which the events occurred. Thus, some children found out about E2 before finding out about E1. If children are merely updating their model of the world sequentially as they find out about each event in the series, their model of the world might erroneously correspond to the circumstances in E2 rather than E1. When the order in which children find out about events differs from the order in which the events actually occurred, and indeed in many other circumstances, updating is insufficient, and instead a *temporal-causal inference* is required. Children must consider the order in which two or more events have occurred and use information about that order to make an inference about a current state of the world. In the case of Povinelli et al.'s task, for instance, children's making the required temporal-causal inference involves them being able to distinguish the order in which E1 and E2 (the hiding events) actually occurred from the order in which they found out about those events and using their knowledge about the former in working out the current location of the puppet.

For present purposes, we define a temporal-causal inference broadly as one in which the causal implications of the temporal order of two or more events must be considered in one's reasoning. There are two obvious circumstances in which temporal-causal reasoning comes into play. A temporal-causal inference could involve taking into account the order of previous events in making a judgment about a current state of the world. This is the type of temporal-causal inference required in Povinelli et al.'s (1999) task. Alternatively, a temporal-causal inference could involve the reverse type of reasoning: using information about a current state of the world to make a retrospective inference about the order in which a series of events must have occurred. In both cases, making the correct inference requires making use of the fact that the causal outcome of the events in the series depends on the order in which the events took place.

Although the distinction between simple updating and temporal-causal inference is a fundamental one, there is little available evidence that would enable us accurately to identify the point in development at which children are capable of making temporal-causal inferences. Indeed, Povinelli's et al.'s (1999) study is the only one that we are aware of that explicitly addresses this issue. The aims of the present studies were thus twofold. First, we wished to examine whether young children were capable of making both types of temporal-causal inferences described above. Can they take into account the order in which two past events occurred to infer a current state of the world, or, conversely, if informed about a current state of the world, can they infer the order in which two past events must have occurred? Two versions of a new task were constructed, one to examine each of these two aspects of temporal-causal reasoning. The second aim was to provide empirical substance to the distinction between temporal updating and temporal-causal reasoning. To fulfill this latter aim, it was necessary to contrast one of the temporal-causal reasoning tasks with a similar task that could be passed by simple updating alone. Thus, across the studies described here, there were three versions of our task in total: two temporal-causal reasoning tasks and a third that could be solved by temporal updating alone.

One way of measuring temporal-causal reasoning is to conduct two (or more) events that the child does not observe and is not aware of, and then to inform the child about the occurrence of those events in an order that differs from the order in which the events actually took place. This is the method used by Povinelli et al. (1999). A potential disadvantage of this method is that it requires using a representational medium, such as videotape, that allows the experimenter to vary the order in which children are made aware of the relevant past events. Another way is to conduct two events that the child is aware are occurring, but to conduct those events out of sight of the child so that the child cannot observe the order in which they occur. The child is then subsequently provided with information that would allow him or her retrospectively to work out the order in which the events occurred. This is the method used in our temporal-causal reasoning tasks, which had the following structure. First, children were introduced to two doll characters, Sally and Katy, who always performed tasks in a certain order: They learned that Sally always went first, and then Katy always went next. Children were then shown a novel piece of apparatus—a large yellow box with two differently colored buttons—and learned how it worked (see Figure 1). Pressing one of the buttons caused a toy car to appear on a shelf in a

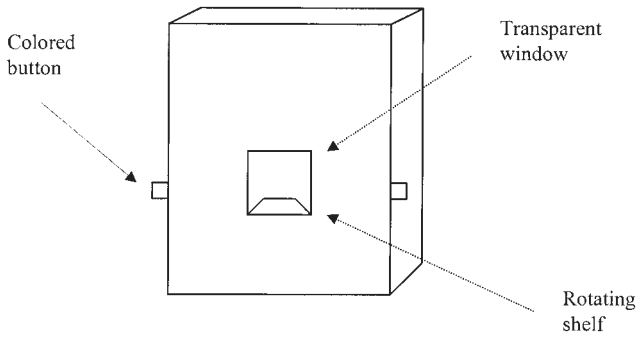


Figure 1. Schematic drawing of the apparatus.

transparent window, whereas pressing the other button caused a marble to appear on the shelf in the window. The window only ever contained one toy at a time, as the box was mechanically constructed such that whatever toy was already in the window dropped out of sight before a new toy appeared (the shelf that held the object rotated automatically whenever a button was pressed to release a new object, taking the old object out of sight). A screen was then put in front of the box and, out of view of the child, Sally and Katy pressed one button each. Finally, the box was uncovered again.

In Study 1, there were two versions of the task testing the two types of temporal-causal inference described above. In the *infer-object* version, after the screen was removed, children could see each doll standing next to the button that she had pressed but the window in the box was left occluded. In this version of the task, children had to infer which toy was inside the window. In other words, given information from which they could determine the order in which two earlier events had occurred (the pressing of the buttons), children had to infer the current state of the world (the object in the window). There was also an *infer-agent* condition. In this version of the task, children were told that Sally and Katy had pressed one button each, but the dolls did not stand next to the buttons they had pressed when the screen was removed. Instead, children were shown the contents of the window and were asked to place each doll beside the button that she had pressed. That is, given information about the current state of the world, children had to infer the order in which two earlier events had occurred. Both of these tasks require making a temporal-causal inference as we have defined it, insofar as correct performance depends on considering the causal implications of the temporal order of the two button-pressing events.

To examine the distinction between temporal updating and temporal-causal reasoning, we used a third version of the task (Study 2). The *visible* version of the task was identical to the *infer-object* task, except for one important procedural change during the testing phase. During this phase, children were actually shown the dolls sequentially pressing their buttons rather than the buttons being pressed behind a screen. Children were still unable to see what was actually in the window following each button press, and after the two buttons had been pressed they were required to judge the contents of the window. Allowing the children to see each button press as it occurred should enable children to sequentially update their model of the contents of the window,

thus we would predict that even quite young children should pass this version of the task.

The *infer-object* and *infer-agent* tasks are in some ways similar to the task used by Povinelli et al. (1999), in that two events occur out of sight and children subsequently do not find out about these events in the order in which those events have occurred. Thus, simple updating is insufficient to pass these tasks. Children do not see the button-pressing events as they happen; rather, they must use the information they are subsequently provided with to either infer a current state of the world (*infer-object* task) or to figure out the order in which the events happened (*infer-agent* task). However, although our *infer-object* task has a similar requirement to that of Povinelli et al., in that it requires using order information to infer a current state of the world, it obviously differs in many other ways. One important difference, mentioned earlier, is that the task does not involve using videotaped events. Although Povinelli et al. argued convincingly that young children's difficulties in their study do not stem from the use of videotapes, it is important to establish whether young children have global problems with temporal-causal reasoning that extend to circumstances that do not involve representational media such as videotapes or photographs.

If we are correct that our temporal-causal reasoning tasks tap a similar ability to Povinelli et al.'s (1999) task, we can make predictions about performance based on their findings. We should expect children who are age 5 to be capable of passing these tasks, whereas children below this age should have difficulties. In fact, Povinelli et al.'s studies found that 5-year-olds were only significantly above chance when the experimenter specifically emphasized the order in which the two causally relevant events occurred. Therefore, we might expect 5-year-olds to need similar cues. With regard to the visible version of the task, which we argue only assesses temporal updating abilities, we would predict that children younger than 5 should find this task straightforward.

Study 1

In this study, children were given either the *infer-object* or the *infer-agent* task. It was not clear in advance of the study whether one of these tasks would prove to be easier than the other. On the one hand, it might be possible to argue that the *infer-agent* condition should be harder, insofar as working out the order of the button-pressing events might involve in some sense "mentally undoing" the current outcome (e.g., reasoning that *although there is a car in the window now*, first there was a marble [so Sally should go by the red button] and only after that did the car appear [so Katy should go by the blue button]). Thus, it might be possible to argue that the *infer-agent* condition requires a certain degree of reversibility of thought, in the Piagetian sense. Arguably, the *infer-object* condition does not require such mental undoing of a current outcome, and children could potentially pass the *infer-object* condition by reasoning through the sequence of events that happened in the order that they happened (e.g., Sally pressed the red button, giving a car, then Katy pressed the blue button, giving a marble, therefore there is marble in the window now). On the other hand, it might be argued that the *infer-agent* condition should be easier, because by its very nature it requires children to focus on the dolls and on the order in which they pressed the buttons. In the *infer-object* condition, the task cannot be passed unless children

focus on the crucial information provided about the order in which the buttons were pressed, which is given by the location of each doll. Potentially, children might ignore the dolls and their locations and try to guess the contents of the window without grasping that the information is available to them to infer the correct answer. In the infer-agent condition, children are handed the dolls and required to put them in the correct locations, ensuring that they focus on the dolls and their locations. On this analysis, the infer-agent condition might be predicted to be easier.

Method

Participants

Forty-five four-year-olds (mean age = 56 months, range = 51–59 months) and 49 five-year-olds (mean age = 65 months, range = 60–71 months) took part in the study. There were 54 boys and 40 girls. Three additional children were tested who were subsequently excluded from the sample, 1 child because he had a formal diagnosis of autism and 2 children because they proved too distractible for the testing protocol to be followed correctly. Twenty-two four-year olds and 26 five-year-olds took part in the infer-object condition, and 23 four-year-olds and 23 five-year-olds took part in the infer-agent condition. Children were recruited from primary schools and nursery schools local to our universities. Eighty-four of the children were of Caucasian descent, 4 were of Afro-Caribbean descent, and 7 were of Asian descent. Children were primarily from working- and middle-class backgrounds. They were tested individually in a quiet corner of their schools, and each child received a set of stickers of his or her choice for taking part.

Materials

Two clothed wooden dolls were used that differed in the color of their clothes and of their hair. Sally was 11 cm high, and Katy was 8.5 cm high. Three different sets of toy props were used in the training phase of the study. There was a plastic shop counter, on which there were two miniature toys of appropriate size for the dolls; one was a toy motorcycle and the other a toy train. On a second shop counter, there were two stickers, one depicting a flower and the other depicting a smiling face. There were also two toy dogs, one white and one black, which were placed next to a set of miniature plastic plants.

A large, electrically powered yellow box apparatus was used in the test phase of the experiment (see Figure 1). The box was approximately 50 cm high, 44 cm wide, and 26 cm deep. On one side of the box there was a red button and on the other side there was a blue button, with each button located 11 cm from the base. In the middle of the front of the box there was a window 8 cm × 8.5 cm that was covered by a hinged transparent plastic window. Inside the box (out of sight of the participants) were two rotating hoppers that could store up to five objects in separate chambers. Each hopper was operated by one of the buttons on the side of the box. When a button was pressed, the hopper rotated so that one of the chambers moved above the entrance of a chute, and any object in the chamber would then drop down the chute into view on a platform located behind the transparent window in the front of the box. The platform onto which objects fell was linked to the rotating mechanism of the hoppers, such that when a button was pressed, both the platform and the hopper rotated in synchrony. When the platform rotated, any object that was already sitting on it fell away into a drawer at the bottom of the box, to be replaced by the new object from the hopper. Thus, there was only ever a single object in the window at any one time, with each old object being replaced by a new one when a button was pressed. When the hoppers and platform were rotating, a mechanical noise could be heard that was identical for each button and could not be localized to one or other side of the box. The objects dispensed by the box

were small marbles of different sizes and colors, and miniature toy cars of various types and colors.

Procedure

Training phase (i). Children were introduced to Sally and Katy and were told that Sally was older than Katy. They then watched Sally and Katy engage in three activities with different sets of props. In the first activity, Sally and Katy visited the toy shop, where there were two toys on the shelf. Children were told that Sally and Katy would get one toy each and that Sally would go first because “Sally always goes first.” Sally removed one of the toys, and then the children were told to watch as Katy went next and removed the other toy, because “Katy always goes next, after Sally.” The other two activities were identical in structure and accompanied by the same wording; in the second activity the dolls chose a dog each to play with, and in the third activity the dolls chose one of two stickers each. After the children viewed the dolls engaged in all three activities, the dolls were set aside and the children were told that they would return to them later.

Training phase (ii). In this phase, children were introduced to how the yellow box worked. The experimenter pointed out the two colored buttons on either side of the box and urged the child to watch what happened when one of the buttons was pressed. The experimenter then pressed either the red or the blue button, and then either a toy car or a marble appeared in the window. Which button was pressed and which button yielded which toy were counterbalanced. The experimenter removed the object, showed it to the child, and asked the child what it was. This was to obtain the child’s preferred name for the marble, because it had been established in pilot testing that not all children were familiar with the word *marble*. The majority of children spontaneously referred to the marble as a *little ball*; all children used the word *car* to describe the toy car. The experimenter replaced the object in the window and then pressed the same button again. The child watched as the object disappeared to be replaced by another object of a different color but of the same type. The experimenter provided a verbal description of what was happening, for example, “Look, that car goes away. And now look, we get a different car.” The children were then encouraged to guess what would appear in the window if the experimenter pressed the same button again, with the experimenter providing the suggestion if the child did not spontaneously name the object type. The experimenter then pressed the same button a third time, and after the new object appeared in the window it was removed and set aside. The experimenter then repeated the same procedure with the other colored button.

The first pair of pretest questions was then asked. Children were asked, “If I press this red [blue] button, what will we get here in the window?” The order in which each button was asked about was approximately counterbalanced between children (43 children were asked about the red button first, and 51 were asked about the blue button first). If the child got either one of the pretest questions incorrect, the experimenter repeated the training procedure involving the box, pressing each of the buttons three times again. If the child again failed either one of the pretest questions, the procedure was halted and the child was not included in the testing. Following the first pair of pretest questions, the dolls were reintroduced to the child, and the child was told that the dolls were going to play with the yellow box and that they were going to press one button each. Children were then asked, “You remember Sally and Katy! Let’s think about them. Which one of them always goes first?” Children were allowed to either name one of the dolls or point to one of them; none of the children got this question wrong. They were then asked the second pair of pretest questions. Their attention was drawn to the two buttons on the yellow box, and the experimenter said, “If we wanted to get a toy car [marble (little ball)], which button would we press?” Once the child had answered, the question was repeated for the other object. The order in which the children were asked about each object was approximately counterbalanced (53 children were asked about the car first, and 39 children were asked about the marble first). If the children got either of these questions incorrect, the training phase involving the box was repeated from the start. If the children again

got one of the second pair of pretest questions wrong following this extra training, the procedure was terminated and the children were not tested.

Test phase. Children were told that Sally and Katy were going to play with the box, and were going to press one of the buttons each, but that they would not be able to see Sally and Katy pressing their buttons. In the infer-object condition, the window of the box was first covered with a small square of card that slotted just behind the transparent cover. The entire front of the box apparatus was then covered with a folding purple cardboard screen. The two dolls were taken behind the screen, and the experimenter pressed first one button and then the other button. Children could hear the (identical) mechanical noises that resulted from each button being pressed but could not see the buttons being pressed. The experimenter placed each one of the dolls beside one of the buttons, with the location of the dolls counterbalanced across children. The large screen was then removed from in front of the apparatus, and the child could see the dolls each standing beside one button but could not see what was inside the window as it remained covered with the card. Children were reminded that the dolls had pressed one button each and that the dolls were standing beside the buttons that they had pressed. The experimenter then said, "OK, now you know which button each doll pressed. What do you think is inside the window?"

In the infer-agent condition, the procedure was similar except that the window was never covered with a piece of card. As in the infer-object condition, the experimenter pressed first one button, and then the other, while the apparatus was covered with the large purple screen. When the large screen was removed, the experimenter said, "Now, you can see what is inside the window. Yes, it's a car [marble]" (counterbalanced between children). The experimenter handed the child the two dolls and said, "OK, I want you to work out which of the dolls pressed which button. I want you to put each doll beside the button that you think she pressed."

Results and Discussion

Two of the 4-year-olds who took part in the infer-object task failed the pretesting questions twice and thus were not tested. Ten of the 4-year-olds and 7 of the 5-year-olds required repetition of the second training phase. Table 1 shows the percentages of children in each group who answered the test question correctly. The 4-year-old group did not perform at a level exceeding chance on either the infer-object task (binomial test, $p = .82$) or the infer-agent task (binomial test, $p = .41$). The 5-year-olds did better on both tasks, with the majority of 5-year-olds passing the infer-object task and the infer-agent task. However, the 5-year-old group did not perform at a level significantly above chance on the infer-object task (binomial test, $p = .33$) and did not significantly outperform the 4-year-old group on this task. On the infer-agent task, the 5-year-olds were performing at a level significantly above chance (binomial test, $p < .05$) and performed significantly better than the 4-year-olds (Fisher's exact test, two-tailed, $p < .05$).

In summary, 4-year-olds did not pass either the infer-object or the infer-agent task, and 5-year-olds' performance was signifi-

cantly above chance only on the infer-agent task. Our findings are thus broadly consistent with those of Povinelli et al. (1999, Study 5), despite the differences between our task and the one used by Povinelli et al. They also found that a group of children younger than 5 had difficulties making an inference about a current state of the world by exploiting information about the order in which two events had happened. They found that even a 5-year-old group had some difficulties with this task, although they found that 5-year-olds did better if the order of the past events was emphasized at testing (Povinelli et al., 1999, Study 6). The finding that the 5-year-old group was significantly better than the 4-year-olds and significantly above chance on the infer-agent task suggests that children of this age do have some grasp of the causal relevance of the temporal order of events. However, we need to consider why this group performed above chance on the infer-agent task but not on the infer-object task.

On our analysis, passing infer-object task requires (a) attending to the locations of the dolls, (b) retrieving the information acquired during training about the order in which the dolls act, (c) retrieving the information acquired during training about which button gives which object, and (d) putting together these two pieces of information. Children's difficulties could potentially result from any one of these four components of the task. Arguably, components (a)–(c) are somewhat easier in the infer-agent condition than in the infer-object task. First, the test question in the infer-agent condition immediately requires children to focus on the dolls themselves. Second, the nature of the question and the focusing of the child's attention in this way may also act as a strong cue for retrieving the information acquired in training about the order in which the dolls always act. More tentatively, a third possibility is that actually being able to see one of the two toys in the window that they encountered in training acts as a good retrieval cue for the information acquired in training about the object yielded by each button. In Studies 3 and 4 we altered the test question in the infer-object task in an attempt to make components (a) and (b) of the task easier, by getting the children to focus on the dolls and the order in which they always do things.

Study 2

The results of Study 1 indicate that 4-year-olds have difficulty making temporal-causal inferences in which they are required to reason about the order in which two events occurred, and that even 5-year-olds may still have some difficulty making such inferences. We have contrasted this type of reasoning with simple updating, in which children need only update or change their model of the world as a series of two events or more unfolds. We have suggested that very young children are capable of updating, providing they can keep track of the events as they occur or are told about the events in the order in which they happened. If this is correct, then even children of 4 years and perhaps younger should be able to pass a version of the infer-object task in which they do not need to retrospectively consider the order in which the dolls pressed the buttons but rather view each button-pressing event as it occurs. To examine this prediction, we tested 3-year-olds on the visible version of the infer-object task, in which the large purple screen was not placed in front of the box apparatus as the dolls pressed the buttons. All other aspects of the task stayed the same.

Table 1
Percentage of Children in Each Age Group Who Answered the Test Question Correctly in Studies 1, 3, and 4

Study and condition	4-year-olds	5-year-olds
Study 1, infer agent	39	74
Study 1, infer object	55	62
Study 3, infer object	55	56
Study 4, infer object	31	71

In addition to examining whether 3-year-olds are capable of accurate updating on this task, this modification also potentially allows us to examine whether older children's problems stem from difficulties understanding how the apparatus works or remembering which button gives which object. The modified version of the task places many of the same requirements on the child as the original version, except it does not require a retrospective temporal-causal inference to be made. Thus, if 3-year-olds perform successfully on the modified task, it would suggest that older children's difficulties lie with making the appropriate temporal-causal inference.

Method

Participants

Participants were 22 three-year-olds (mean age = 44 months, range = 39–50 months; 13 boys and 9 girls). Twenty of the children were of Caucasian descent, and 2 were of Afro-Caribbean descent. Children were drawn from a nursery for the children of staff and students at Queen's University and were tested individually in a quiet room in the nursery. Children received a set of stickers of their choice for participating in the task.

Materials

These were identical to those used in Study 1.

Procedure

The training procedure was identical to that of Experiment 1. The testing phase of the experiment was also identical to the infer-object condition of Experiment 1, except that the large purple screen was not placed in front of the apparatus while the dolls pressed their buttons. Instead, only the transparent window was covered, and children watched as first Sally and then Katy pressed one of the buttons. Children were asked to infer which object was in the window after they had seen both buttons being pressed.

Results and Discussion

Four children who failed to complete training were excluded from the test phase; none of the remaining children required repetition of any of the training phases. Of the remaining 18 children, 15 correctly judged which object was behind the window. Thus, children were significantly above chance (binomial test, $p < .01$) on the visible version of the infer-object task. These results contrast with the findings of Experiment 1, in which even 5-year-olds had difficulty with the infer-object task. As we have emphasized, the only difference between the two versions of the task is whether children could actually see the two buttons pressed one at a time by the doll characters, as opposed to finding out shortly afterward which doll pressed which button. Our findings provide support for our distinction between updating and making genuine temporal-causal inferences, with 3-year-olds being competent at the former.

Study 3

The good performance of 3-year-olds on the visible version of the infer-object task is in contrast to older children's at-chance performance on the original version of the task and again raises the question of what exactly children find so difficult about this task.

We have suggested that perhaps children may simply guess what object is inside the window, rather than using the crucial information provided by the position of the two dolls about the order in which the buttons have been pressed. This may be because children either fail to attend to the locations of the dolls or fail to retrieve or utilize their knowledge acquired at training. Perhaps young children are capable of utilizing such information if their attention is explicitly drawn to it and they are encouraged to think about it. This possibility was examined in Study 3, in which children were given a cue to attend to the dolls and to think about the order in which the dolls act before answering the test question.

Method

Participants

Thirty-one four-year-olds (mean age = 56 months, range = 48–59 months) and 32 five-year-olds (mean age = 65 months, range = 60–71 months) from primary schools and nursery schools local to our universities took part in the study. There were 41 boys and 22 girls. One additional 4-year-old girl who was tested was not included in the sample because she would not supply an answer to the test question. All of the children were of Caucasian descent, and they were primarily of middle-class background. None of the children had previously taken part in Study 1. Children were tested individually in a quiet room in their schools and received a set of stickers for their participation.

Materials

These were identical to those used in Studies 1 and 2.

Procedure

The training and test procedures were identical to those used in the infer-object condition in Study 1, except for a single modification. At the test phase, the large purple screen was removed and the experimenter said:

OK, I want you to think carefully. Which doll went first? That's right [children never answered this question incorrectly]. So Sally went *first* and pressed her button and then Katy went *next* and pressed her button. Look at the buttons Sally and Katy pressed. There is one thing inside the window now. What do you think is inside the window right now?

Thus, children's attention was explicitly drawn to the dolls and to the order in which the buttons were pressed.

Results and Discussion

None of the children failed to complete training. The second phase of training had to be repeated for 6 of the 4-year-olds. Table 1 shows the percentage of children of each age group who answered the test question correctly in this modified version of the infer-object task. It can be seen that these percentages are very similar to those in Study 1, despite the modification of the test instructions. Again, a marginal majority of 5-year-olds did answer correctly, but as a group they were not above chance (binomial test, $p = .60$), and the 4-year-old group was also not above chance on this task (binomial test, $p = .72$). The difference between the two groups was also not significant (Fisher's exact test, $p > .05$).

Thus, despite drawing children's attention to the dolls and to the order in which the button-pressing events occurred, performance was still not very good on this task. Performance on this task can

be contrasted with the excellent performance of 3-year-olds in Study 2, in which the actual button-pressing events could be observed. In the present task, although the button-pressing was hidden from view, children's attention was drawn by the experimenter to the dolls and to the order in which the buttons had been pressed, with the experimenter restating what had happened out of view. Povinelli et al. (1999) have argued that children below age 5 do not grasp the significance of order information in coming to a conclusion about the current state of the world. This may be true for our younger children too; however, given Povinelli et al.'s (1999, Study 6) findings, and the findings of the infer-agent task in Study 1, it is difficult to conclude that 5-year-olds' problems with the infer-object task stem from a lack of this basic type of understanding.

One possible reason for the poor performance of the 5-year-olds may be that the task protocol makes one of the doll characters—Sally, who goes first—more salient than the other. When children were questioned about order (which occurred twice in the protocol for Study 3), they were asked, "Which doll goes first?" Making Sally more salient in this way may lead to difficulties at test, in that although she pressed a button first, it is the button that Katy pressed that determines which object is in the window when children are asked the test question. The task procedure was altered in Study 4 to try to balance the salience of the two doll characters.

Study 4

Method

Participants

Twenty-six four-year-olds (mean age = 56 months, range = 52–59 months) and 35 five-year-olds (mean age = 65 months, range = 60–69 months) took part in the study, from primary schools and nursery schools local to our universities. There were 37 boys and 24 girls. None of the children had participated in any of the other studies. Two additional 5-year-olds were tested but were not included in the sample because the task protocol was not followed correctly (both children removed the screen from the window before answering the test question). All of the children were of Caucasian descent, and they were primarily of lower- to middle-class background. Children were tested individually in a quiet room in their schools and received a set of stickers for their participation.

Materials

These were identical to those used in the previous studies.

Procedure

The procedure was identical to that used in Study 3, except for three alterations in the training phase and test instructions. First, during the first phase of training, what was explicitly emphasized was not which doll went first, but which doll went last ("Remember, Katy always goes last"). Second, in the pretest questions, participants were asked, "Which doll always goes last?" rather than which doll went first. Third, after the purple screen was removed at testing, participants were asked, "So which doll pressed her button last?" and after they had answered this question, the experimenter said, "OK, so Sally pressed her button and then Katy pressed her button. Look at the buttons that the dolls pressed. There is one thing inside the window right now. What do you think it is?"

Results and Discussion

None of the children failed to complete training. Sections of the training phases were repeated for 5 of the 4-year-olds and 1 of the

5-year-olds. Table 1 shows the percentage of children of each age group who gave the correct answer to the test question. The 4-year-olds performed poorly on this task, with their performance not differing significantly from chance (binomial test, $p > .05$). However, the performance of the 5-year-olds was significantly above chance (binomial test, $p < .02$). Their performance was significantly better than that of the 4-year-olds (Fisher's exact test, $p < .01$).

The alteration of the task procedure to change the salience of the doll that always goes first would seem to have been successful in improving 5-year-olds' performance on the task. The findings suggest that the failure of the 5-year-olds in the infer-agent task in Studies 1 and 3 was not due to a failure to grasp the significance of temporal order per se but to focus on the correct piece of causally significant information. There are in fact two, related, possible descriptions of 5-year-olds' difficulty in Study 3. Perhaps, as suggested, in the previous studies children found Sally more salient and based their answer simply on what Sally had done, rather than shifting attention to what Katy, the doll who went last, had done. A second, related, possibility is that the use of the word *first* in itself is problematic for children of this age (see also Piaget, 1969). Although, in this context, *first* is clearly intended to be of purely temporal significance, it is possible that the children had a wider reading of the term than this, interpreting it as meaning the most important doll—the one who has priority or who gets what she wants. Given that Sally was also the larger doll, and introduced as older than Katy, it seems possible that even the 5-year-olds interpreted *first* in this way. It may be that simply removing the term *first* from the task narrative helped improve the performance of this group.

General Discussion

The aim of these studies was to establish a new task to examine whether young children can reason about the causal significance of the temporal order in which past events have occurred, as opposed to simply being able to update their model of the world as changes unfold. One important difference between the present paradigm and that used by Povinelli et al. (1999) is that our task does not require use of a representational medium such as videotape that young children may have difficulty with. Despite using a very different paradigm, our results closely resemble those of Povinelli et al. (1999). Children below age 5 had difficulties making temporal-causal inferences, even under circumstances in which they were cued to attend to the relevant temporal order information (Studies 1, 3, and 4). Five-year-olds were capable of making temporal-causal inferences under some circumstances. In Study 1, they passed the infer-agent version of the task, which involved inferring the order in which two events must have happened, given a current state of the world. In Study 4, they also passed an infer-object version of the task, in which they were cued to use temporal order information to infer a current state of the world.

Three-year-olds were successful on a visible version of the infer-object task. This task was identical to the task used with older children in the other studies apart from the fact that, in the visible version, the children were actually able to see the dolls pressing their buttons (rather than being shown the dolls standing beside the buttons that they had just pressed, as in the normal version of the task). Although this alteration of the task procedure may seem on

first consideration to be slight, we would argue that it is important because it enabled children to use updating to pass the task rather than making a temporal–causal inference. Successful performance simply required changing one’s model of the world (i.e., of the contents of the window) sequentially as each button was pressed, rather than having to recruit information about the temporal order in which the buttons were pressed to retrospectively infer a state of the world. The successful performance of the 3-year-old group on the visible version of the task provides support for making a developmental distinction between updating and temporal–causal inference. Although intuitively plausible, there is little previous empirical evidence that we are aware of that provides a basis for this specific distinction.

Our findings provide further evidence that by at least 5 years, children are capable of making temporal–causal inferences. However, even 5-year-olds seemed to have difficulty in some task conditions. In the infer-object task in Study 1, children had to spontaneously recruit order information that was implicit in the display in making their inference about a current state of the world. Five-year-olds were not above chance on this task, consistent with the findings of Povinelli et al. (1999), who also found that a 5-year-old group was not significantly above chance on a version of their task in which children were not cued to use temporal order information. They found, however, that this age group performed above chance when encouraged to think about the order in which the previous events had occurred. Given Povinelli et al.’s findings, we hypothesized that 5-year-olds’ performance might improve if encouraged to attend to and use the information provided in the display in the infer-object task that would allow them to make the correct inference. This hypothesis is also suggested by the fact that this age group was above chance in the infer-agent task. We have suggested earlier that the questioning in the infer-agent task by its nature requires children to attend to the dolls and to think about information acquired in training about the order in which the dolls always act. It is possible that in the infer-object task, children fail to attend to the information provided to them in the display and simply guess the answer. Therefore, we predicted that 5-year-olds should perform above chance on this task if the relevant information on which to base their reasoning is made explicit to them.

However, contrary to this prediction, providing this sort of cue did not help even this older group (Study 3). The findings of Study 4, in which 5-year-olds did perform successfully, suggest that the difficulties of this age group on the infer-object task in Studies 1 and 3 may have stemmed from the exact nature of the task instructions and temporal cue that they were provided with. The instructions used in Studies 1 and 3 made the doll that “always goes first” the most salient doll and may have led to an underestimation of 5-year-olds’ abilities, because it is the actions of the doll that went last that were in fact crucial in determining the answer to the task question. In Study 4, in which children were encouraged to attend to the doll that always went last, 5-year-olds performed above chance.

In fact, this latter finding raises the possibility of an alternative analysis of the demands of the task. We have suggested that the infer-object task is a temporal–causal reasoning task insofar as it involves reasoning about the causal consequences of the order in which the two button-pressing events took place. However, it is possible to argue that the infer-object task, as it stands, could be passed simply by attending to one of the dolls, the doll that went

last, without considering both button-pressing events and the order in which they occurred.¹ This interpretation is made possible because of the nature of the causal relationships used in the studies. Specifically, the outcome of pressing each button is unaffected by the sequence of events that precede it: In any given testing session, the act of pressing a red button, for example, always yields a car, regardless of what buttons have been pressed before it. An alternative scenario would have been to construct the apparatus such that preceding events did have causal consequences for the outcome of subsequent events: for example, designing the apparatus such that pressing the red button only yielded a car if the blue button had been pressed prior to it. Thus, it may be possible to argue that children need only focus on the last button-pressing event, ignoring the previous button-pressing event and the order in which the two button presses took place. The fact that 5-year-olds pass the infer-object task when the act of the last doll is emphasized might be thought to lend support to this interpretation, insofar as the task instructions focus children’s attention on this doll. There are two possible responses to such an argument. One response may be to argue that grasping that the last act of two acts is the causally relevant one in itself involves some sort of consideration of the relative order of the acts—the last act is only such because there were one or more preceding acts. Another response might be that if passing the task is a matter of simply focusing on the actions of only one doll, then the 4-year-olds might be expected to perform well in Study 4, in which the last doll was especially salient. In fact, the performance of this group was at its lowest level in this study. Nevertheless, we acknowledge that a possible fruitful avenue for future research would be to vary the nature of the causal relationships between events in a sequence. Specifically, performance on the present task could be contrasted with performance on a task in which the outcome of the last event in the sequence depended causally on what events had preceded it.

We have argued that, taken as a whole, our findings indicate a developmental progression from reliance on temporal updating to the ability to make temporal–causal inferences. However, it should be noted that there are a variety of reasons why the younger children may fail to make the appropriate temporal–causal inference. The pretest questions were designed to ensure that all children who progressed to the testing stage knew (a) the order in which the dolls always acted and (b) which toy was yielded by which button. We are confident that children knew the order in which the dolls acted, insofar as no child ever failed the relevant pretest question. Indeed, in Studies 3 and 4 they were questioned about this again immediately before the test question, and again no child got this question wrong. We would also argue that the good performance of the 3-year-olds on the visible version of the infer-object task suggests robust memory for how the apparatus worked, even in these very young children. Nevertheless, as a further check, memory for this information could have been reassessed posttest.

Putting this aside, if we make the assumption that children’s difficulty lay not in remembering the two types of relevant information but in putting them together, there are at least two alternative explanations for our findings. First, it may be that young

¹ We are grateful to an anonymous reviewer for suggesting this interpretation.

children do not understand the significance of the temporal order information they have been provided with. In other words, the failure is a conceptual one: As Povinelli et al. (1999) described it, children of this age do not grasp that the passage of time involves the unfolding of “a successive series of causally interdependent states of the world” (p. 1427). Second, it may be that children have the relevant concept of time, but their difficulties stem from the working memory demands of the task. The idea here would be that it is the combined processing demands of retrieving the necessary pieces of information from long-term memory, holding them in working memory, and making the appropriate inference that are beyond younger children. In this interpretation, children’s failure is not due to a lack of understanding, but of putting that understanding to work. Our findings, as they stand, do not allow us to distinguish clearly between these two possible interpretations. The fact that 4-year-olds are still not successful in Study 4, when they are provided with some assistance in the reasoning process, would seem to point toward a failure of understanding. However, this issue may be best resolved by examining children’s performance on a variety of different temporal–causal reasoning tasks that differ in their working memory demands. We are currently pursuing this line of investigation.

Finally, we turn to considering some of the broader theoretical implications of our findings. If it is correct that our task and that of Povinelli et al. (1999) shed light on a conceptual limitation in young children, this has important implications for descriptions of children’s developing understanding of time. Historically, research on the development of children’s thinking about time has tended to focus on the emergence of more complex temporal abilities (Fraissee, 1982; Piaget, 1969), perhaps due to a lack of useful analyses of the components of basic temporal understanding (for notable exceptions, see Nelson, 1996; Weist, 1989). However, although Piaget focused on relatively complex temporal abilities, he also made it clear that an important component of a mature concept of time is that identified by Povinelli et al. (1999), namely an appreciation of the systematic causal relationships that obtain between events that are associated with their temporal order (Piaget, 1969, especially chap. 1). Furthermore, the idea that a mature or objective conception of time involves a certain type of causal understanding is one that has appeared previously in the philosophical literature (Campbell, 1994; Martin, 2001).

One important set of considerations, in this context, concerns potential connections between a grasp of the asymmetrical nature of causal relations and an understanding of the notions of the past and the present. At its simplest, the idea would be that understanding the notions of the past and the present is, in part, a matter of understanding that what was the case in the past may have had a causal impact on what is the case in the present, but that the converse is not possible. However, an equally crucial ingredient in our mature notion of time is the idea that certain past events have, or might have, happened without any causal traces of them remaining in the present. (In philosophy, this issue is closely connected to debates about realism concerning the past; see, e.g., Peacocke, 2001.) It is this aspect of temporal thought, specifically, that seems particularly closely connected to the specific type of temporal–causal reasoning we have described in this article. That is to say, what makes intelligible the idea that certain past events might have happened without leaving any causal traces in the present is precisely the thought that the overall outcome of a

sequence of events depends on the temporal order in which they happen, such that later events in the sequence may change or obliterate the effects of events that occurred earlier in the sequence.

This is not to say that young children who may lack such causal understanding of event sequences have no way of conceptualizing the relationships between different events. The fact that children use tensed forms of verbs, often correctly, early in their linguistic development (i.e., at an age younger than the children tested in the present studies) suggests otherwise (Smith, 1980; Weist, 1986, 1989). However, the challenge for cognitive developmental psychologists is to provide sufficiently rich descriptions of early conceptual abilities that may underpin this and other related abilities, and to describe how these differ from the possession of mature temporal concepts (McCormack & Hoerl, 1999). The suggestion of the present article is that providing such a description must at least in part rest on a consideration of how causal reasoning abilities develop and change.

In conclusion, the findings of the present studies provide support for a developmental distinction between the ability to update one’s model of the world as it changes and more sophisticated temporal–causal reasoning abilities. The ability to make temporal–causal inferences is likely to be one of considerable developmental importance, the emergence of which enables children to deal more effectively with temporally extended reality. In particular, a number of researchers have argued that temporal–causal understanding is important in the development of autobiographical or episodic memory (McCormack & Hoerl, 2001; Perner, 2001; Pillemer, 1998; Povinelli et al., 1999; Reese, 2002; Welch-Ross, 2001). Further, although we have only considered here children’s ability to make temporal–causal inferences about events that have already occurred, this ability may also be vital in planning events in the future. Indeed, recent research would suggest that preschool children have difficulties planning future sequences of events, even if they can remember and reconstruct similar sequences of events (Atance & O’Neill, 2001; Benson, 1997; but see Bauer, Schwade, Wewerka, & Delaney, 1999). Such difficulties may stem from problems in explicitly reasoning about the causal connections between events. Given the fundamental nature of such reasoning, it would seem to be important that future research examines whether the developmental findings that we have reported here are generalizable and that it explores the developmental role of temporal–causal reasoning in a variety of cognitive domains.

References

- Atance, C., & O’Neill, D. K. (2001). Planning in 3-year-olds: A reflection of the future self? In C. Moore & K. Lemmon (Eds.), *The self in time* (pp. 121–140). Hillsdale, NJ: Erlbaum.
- Bauer, P. J., Schwade, J. A., Wewerka, S. S., & Delaney, K. (1999). Planning ahead: Goal-directed problem solving by 2-year-olds. *Developmental Psychology*, 35, 1321–1337.
- Benson, J. B. (1997). The development of planning: It’s about time. In S. L. Friedman & E. K. Scholnick (Eds.), *The developmental psychology of planning: Why, how, and when do we plan?* (pp. 43–76). Mahwah, NJ: Erlbaum.
- Bullock, M., Gelman, R., & Baillargeon, R. (1982). The development of causal reasoning. In W. Friedman (Ed.), *The developmental psychology of time* (pp. 209–254). New York: Academic Press.
- Campbell, J. (1994). *Past, space, and self*. Cambridge, MA: MIT Press.

- Corrigan, R., & Denton, P. (1996). Causal understanding as a developmental primitive. *Developmental Review, 16*, 162–202.
- Fraisse, P. (1982). The adaptation of the child to time. In W. Friedman (Ed.), *The developmental psychology of time* (pp. 113–140). New York: Academic Press.
- Gelman, R., Bullock, M., & Meck, E. (1980). Preschoolers' understanding of simple object transformations. *Child Development, 51*, 691–699.
- Gopnik, A., Sobel, D. M., Schulz, L. E., & Glymour C. (2001). Causal learning mechanisms in very young children: Two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental Psychology, 37*, 620–629.
- Leslie, A. M. (1982). The perception of causality in infants. *Perception, 11*, 173–186.
- Martin, M. (2001). Out of the past: Episodic recall as retained acquaintance. In C. Hoerl & T. McCormack (Eds.), *Time and memory: Issues in philosophy and psychology* (pp. 257–284). Oxford, England: Oxford University Press.
- McCormack, T. (2001). Attributing episodic memory to animals and children. In C. Hoerl & T. McCormack (Eds.), *Time and memory: Issues in philosophy and psychology* (pp. 285–314). Oxford, England: Oxford University Press.
- McCormack, T., & Hoerl, C. (1999). Memory and temporal perspective: The role of temporal frameworks in memory development. *Developmental Review, 19*, 154–182.
- McCormack, T., & Hoerl, C. (2001). The child in time: Temporal concepts and self-consciousness in the development of episodic memory. In C. Moore & K. Lemmon (Eds.), *The self in time* (pp. 203–227). Hillsdale, NJ: Erlbaum.
- Nelson, K. (1996). *Language in cognitive development: Emergence of the mediated mind*. Cambridge, England: Cambridge University Press.
- Oakes, L. M. (1994). Development of infants' use of continuity cues in their perception of causality. *Developmental Psychology, 30*, 869–879.
- Oakes, L. M., & Cohen, L. B. (1990). Infant perception of a causal event. *Cognitive Development, 5*, 193–207.
- Peacocke, C. (2001). Understanding the past tense. In C. Hoerl & T. McCormack (Eds.), *Time and memory: Issues in philosophy and psychology* (pp. 257–284). Oxford, England: Oxford University Press.
- Perner, J. (2001). Episodic memory: Essential distinctions and developmental implications. In C. Moore & K. Lemmon (Eds.), *The self in time* (pp. 181–202). Hillsdale, NJ: Erlbaum.
- Piaget, J. (1969). *The child's conception of time*. New York: Basic Books.
- Pillemer, D. B. (1998). *Momentous events, vivid memories*. Cambridge, MA: Harvard University Press.
- Povinelli, D. J., Landry, A. M., Theall, L. A., Clark, B. R., & Castille, C. M. (1999). Development of young children's understanding that the recent past is causally bound to the present. *Developmental Psychology, 35*, 1426–1439.
- Reese, E. (2002). Social factors in the development of autobiographical memory: The state of the art. *Social Development, 11*, 124–142.
- Schlottmann, A., Allen, D., Linderoth, C., & Hesketh, S. (2002). Perceptual causality in children. *Child Development, 73*, 1656–1677.
- Shultz, T. (1982). Rules of causal attribution. *Monographs of the Society for Research in Child Development, 47*(1, Serial No. 194), 1–51.
- Shultz, T., & Kestenbaum, M. (1985). Causal reasoning in children. *Annals of Child Development, 2*, 195–249.
- Smith, C. S. (1980). The acquisition of time talk: Relations between child and adult grammars. *Journal of Child Language, 7*, 263–278.
- Weist, R. M. (1986). Tense and aspect. In P. Fletcher & M. Garman (Eds.), *Language acquisition: Studies in first language development* (2nd ed., pp. 356–374). Cambridge, England: Cambridge University Press.
- Weist, R. M. (1989). Time concepts in language and thought: Filling the Piagetian void from two to five years. In I. Levin & D. Zakay (Eds.), *Time and human cognition: A life-span perspective* (pp. 63–118). Amsterdam, The Netherlands: Elsevier.
- Welch-Ross, M. (2001). Personalizing the temporally extended self: Evaluative self-awareness and the development of autobiographical memory. In C. Moore & K. Lemmon (Eds.), *The self in time* (pp. 97–120). Hillsdale, NJ: Erlbaum.

Received December 10, 2003
 Revision received July 2, 2004
 Accepted July 20, 2004 ■

E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at <http://watson.apa.org/notify/> and you will be notified by e-mail when issues of interest to you become available!