INTERNATIONAL JOURNAL OF BEHAVIORAL DEVELOPMENT, 1990, *13* (1) 95–117

Conceptual and Linguistic Factors in Children's Memory for Causal Expressions

James P. Byrnes and Susan A. Gelman
University of Michigan

The present study examined children's memory for causal expressions as a measure of their implicit knowledge. In experiment I, first-, third-, and fifth-graders were asked to recall descriptions of causal sequences including "if" and "because" statements. Protocols were scored for recall of causal ideas and use of the terms "if" and "because." Results showed significant age differences in recall of causal ideas, but mostly with regard to causal explanations and not concepts of covariance. Significant increases were also found in recall of the terms "if" and "because" to describe these sequences. In experiment II, the same subjects received a sentence recall task that assessed the degree to which they associated "if" with unexpected content and "when" with expected content. On this task, third- and fifth-graders were significantly more likely to associate "if" with unexpected content and "when" with ordinary content. Implications for children's understanding of causal relations and for knowledge-based accounts of memory are discussed.

INTRODUCTION

Recent investigations of memory development have stressed the important role that children's knowledge plays in determining developmental differences in recall (e.g. Brown, Bransford, Ferrara, & Campione, 1983; Chi, 1978; Siegler, 1986). The key role of knowledge has been illustrated in studies showing that children's recall is superior to that of adults if children are more knowledgeable about the to-be-remembered information (e.g. Chi, 1978; Lindberg, 1980). Knowledge has been argued to affect the

Requests for reprints should be sent to either James Byrnes, Institute for Child Study, Department of Human Development, College of Education, University of Maryland, College Park, MD 20742, or Susan Gelman, Department of Psychology, 3433 Mason Hall, University of Michigan, Ann Arbor, MI 48109.

This research was supported by a Spencer Fellowship to the second author. The authors wish to thank two anonymous reviewers for very helpful suggestions.

encoding of information (Siegler, 1980) and the amount of information stored, as well as the nature of retrieval (Flavell, 1985).

In the present study we focused on children's memory for causal expressions. Recent studies have shown development both in knowledge of causal relations (Bullock, Baillargeon, & Gelman, 1982) and in knowledge of certain linguistic aspects of causal expressions (Emerson, 1979, 1980). It is reasonable to expect, then, that significant developmental changes in knowledge would affect recall. In addition, there are certain inconsistencies in the literature that the use of a recall methodology could address.

Causal reasoning has two primary components: attribution and explanation. Causal attribution involves assigning a causal role to antecedent events based on principles such as *covariance*, *temporal priority*, and *generative transmission*. For example, an event is more likely to be deemed causal if it consistently covaries with an outcome (Shultz & Mendelson, 1975), if the event occurs prior to the outcome (Bullock et al., 1982), and if an object within the event appears to transmit something to an object within the outcome (Shultz, 1982). Recent studies have shown that even 3-year-olds use these principles consistently and correctly when making causal attributions. Such knowledge, then, is present quite early and seems to develop little beyond the preschool years. Attribution principles might be considered "cause-general" knowledge since all causal relations conform to these principles.

The second component, causal explanation, refers to the construction of a model that posits a *mechanism* to account for the relationship between cause and effect (Piaget, 1974). That is, after a causal relationship is judged to be likely using principles of attribution, we construct a mental model of just how a particular cause brings about its effect.

The accuracy or reasonableness of mechanisms posed varies with age, cognitive level, and familiarity with the particular mechanism in the causal sequence. For example, even adults often fail to understand the precise mechanism and chain of events involved in causing a car engine to start, but they still believe in the causal role of turning the key. Thus, adults are able to hold a correct belief in the causal role of turning the key without necessarily having an accurate model of the mechanism involved. Given that even adults have difficulty comprehending certain mechanisms, it is not surprising that the only studies which show marked developmental change in knowledge of causal relations require children to give explanations (e.g. Piaget, 1974). However, Berzonsky (1971) showed that varying children's familiarity with a causal sequence greatly reduced the number of inappropriate mechanisms cited by children (e.g. "What makes a bicycle go?" vs "What makes the wind blow?"). Explanations, then, require "cause-specific" knowledge since causes often bring about their effects through particular mechanisms.

In sum, the most marked developmental change in children's knowledge of causal relations concerns the adequacy of explanations (e.g. Piaget, 1974; Berzonsky, 1971). Given the premise that knowledge affects recall, one would expect that developmental differences in recall of causal sequences would be most apparent for explanations, not for attributions.

In addition to knowledge of causal mechanisms, a second possible factor which might affect recall of causal statements concerns children's linguistic knowledge of "if" and "because". Syntactically, children need to come to know the temporal order constraint on "if" and "because" (Emerson, 1979; 1980). This constraint specifies that regardless of whether "if" or "because" appears in the first or second clause of a sentence, whichever event follows the connective happened first, and the event in the other clause happened second. Research regarding children's knowledge of this constraint has found conflicting results. In natural discourse, children as young as age 2 or 3 years produce meaningful "if" and "because" expressions that do not violate the temporal order constraint (Bloom & Capatides, 1987; French & Nelson, 1985; Hood & Bloom, 1979). That is, they rarely say such things as (i) "He fell off his bike because he broke his arm" when they mean (ii) "He broke his arm because he fell off his bike." In experimental contexts, however, 6- to 9-year-olds appear to have little knowledge of this constraint. When asked to judge whether sentences such as (i) and (ii) mean the same thing, children below the age of 10 are often apt to say they do (Emerson, 1979; 1980). This apparent disregard for the temporal order constraint has been called a "reversed-clause error". The inconsistency between spontaneous production and experimental studies, however, can be easily reconciled by noting that children are often able to produce well-formed expressions of a particular kind much earlier than they are able to make metalinguistic judgments about such expressions (Brown, 1973; Slobin, 1979). Metalinguistic judgments, then, are said to underestimate children's syntactic knowledge.

In sum, the primary developmental change in children's syntactic knowledge of causal expressions concerns the elevation of implicit knowledge of the temporal order constraint to explicit, conscious knowledge. It is unclear as to whether such change will affect recall. We empirically address this issue in the present study.

An additional linguistic aspect of causal expressions concerns the pragmatic concepts they convey. In English, a speaker typically uses "if" to convey uncertainty about whether the events described will in fact occur, and "when", "so", or "because" to express certainty about the occurrence of these events (Bates, 1976; Scholnick & Wing, 1982). Compare (iii) "If it rains today, we'll get wet" to (iv) "Because it will rain today, we'll get wet", (v) "When it rains today, we'll get wet", and (vi) "It will rain today so we'll get wet". As was the case for the temporal order constraint,

conflicting results have been obtained. Whereas Bates (1976) has found an appreciation of some pragmatic aspects of "if" in preschoolers, Scholnick & Wing (1982, 1983) found that knowledge that "if" conveys uncertainty and "because" conveys certainty was not apparent until 9–10 years of age. Again though, children were required to make explicit, metalinguistic judgments about someone else's utterances in the latter studies. Hence, development occurs largely with respect to the shift from implicit to explicit knowledge of the pragmatic concepts conveyed. We consider possible effects of such changes in the present study.

Two experiments were conducted in the present study. In experiment I we demonstrated causal sequences, described these sequences with "if" and "because" statements, and asked children to recall these descriptions. Our recall methodology derives from Bransford & Franks (1972) and others reviewed in Paris (1978). We assume that the form of children's recall of causal descriptions provides insight into their conceptual and linguistic knowledge. In particular, if knowledge affects recall, then differential patterns of recall should be consonant with research findings regarding developmental differences in knowledge. Regarding knowledge of causal relations, one would expect different patterns of recall for cause-general aspects (e.g. covariance relations) than for cause-specific aspects (e.g. explanations). Regarding knowledge of linguistic aspects, one would expect developmental differences if one believes that knowledge of temporal order and pragmatic concepts needs to be explicit and available to consciousness.

In experiment II we assessed children's knowledge of the pragmatic concepts conveyed by connectives. Children were presented with sentences constructed with "if" and "when" and were asked to recall these sentences. Whereas "if" often conveys uncertainty about the content described, "when" often conveys certainty. The connectives were paired with semantic content that was either consistent with these presuppositions or contrasted with them. If children are knowledgeable about these presuppositions, they should be more likely to recall sentences in which the semantic content and presuppositions match than when they do not match.

The recall methodologies used in experiments I and II have advantages over other experimental techniques described earlier in that children are not required to make metalinguistic judgments. If children do not adhere strictly to the temporal order constraint or have little knowledge of pragmatic concepts, they should make a high number of "reversed-clause" errors, and at times confuse "if" with connectives which convey certainty (e.g. "when" or "because"). If on the other hand children do possess implicit knowledge of such linguistic aspects of causal expressions, the present methodology would be more likely to reveal this knowledge than prior experimental techniques.

EXPERIMENT I

Method

Subjects. Twenty-six children each from the first (mean age 6–8 years), third (mean age 8–11 years) and fifth grades (mean age 10–18 years) participated. There were 16 boys and 10 girls in the first grade, 12 boys and 14 girls in the third grade, and 11 boys and 15 girls in the fifth grade. All children spoke English as a primary language and were drawn from three public elementary schools in a university town.

Tasks and Procedure

Apparatus Recall Task. For this task, children listened to causal descriptions of a balance and a buzzer apparatus and were asked to recall the descriptions.

The Balance. The balance apparatus consisted of a 35-cm long wooden balance mounted on a 51 × 15 cm board, and a 1.5-volt electric bulb. Subjects were shown that when the left arm of the balance descends, it touches contacts wired to the bulb and makes it light. Upright wooden dowels served as pegs for ring-shaped weights. On the left arm of the balance, there were three dowels positioned 5 cm, 10 cm, and 15 cm to the left of the fulcrum. On the right arm, a weight was glued in place on a dowel which was 5 cm to the right of the fulcrum. Children were shown that an equal weight placed one unit away from the fulcrum on the left arm balanced the apparatus and did not, therefore, activate the light. Weights placed on the dowels either 10 or 15 cm to the left of the fulcrum caused the balance to descend and activate the light. The dowel nearest the fulcrum on the left arm was called "number 1" and the second and third dowels were called "number 2" and "number 3" respectively. To reduce the memory load for these labels, small cards with numbers on them were placed on the table beneath the respective dowel. Before giving the description of the balance, the experimenter warned the child to listen carefully to the description because he or she would be asked to try and say back exactly what the experimenter said. The experimenter then demonstrated how the balance worked as he read the following description to each child:

"This is a balance. It's like a teetertotter. If I put this weight on spot number 1 [places weight], nothing happens—this side doesn't go down. If I put the weight on spot number 2 [places weight], the weight makes this side go down and turn on the light. If I put the weight on spot number 3 [places weight], the weight also makes this side go down and turn on the light. The

light goes on because the weight is far out at number 2 [points], and even farther out at number 3" [points].

This description was read to children twice. As each was mentioned, the experimenter pointed to the respective spots and arms, moved the weights, and activated the light. The experimenter then asked children to retell the description, saying, "Now it's your turn. I want you to say back to me what I said to you. I want you to try and use the same words I used." The balance was then placed in front of the subject and he or she was instructed to operate it while giving the description.

Scoring. Each subject's recall protocol was scored with respect to conceptual content and linguistic form considered separately. With respect to conceptual content, the balance description was broken down into four causal ideas: (a) spot 1 fails to produce the light; (b) spot 2 produces the light; (c) spot 3 produces the light; and (d) the light is caused by the distance of the weight from the fulcrum. Subjects could receive, then, a score of 0–4 for content. Ideas (a), (b) and (c) are referred to in subsequent analyses as the *covariance information*, and (d) is referred to as the *explanation information*. With respect to linguistic form, the protocols were scored according to whether subjects used "if" when referring to the covariance information and "because" when explaining the causal relations. Hence, each subject could receive a score of 0–3 for "if" and 0–1 for "because". Content and form were scored separately because subjects could recall all four content ideas without using the same linguistic form as the experimenter (although they were encouraged to do so), or could use the same linguistic form as the experimenter but describe the apparatus incorrectly or incompletely. Finally, each protocol was also scored with respect to the frequency of "if"-to-"when" conversions during recall, that is, using "when" to describe the covariance information even though they heard the experimenter use "if". Subjects could obtain a score of 0–3 for conversions.

The Buzzer. The buzzer apparatus consisted of a small 18 × 10 × 8-cm plastic box. Five switches were mounted in a line along the near edge of the top of the box. The leftmost and rightmost switches were painted white and the middle three were black. A small electric buzzer was also mounted in the centre of the far edge of the top of the box and was encased in a 2-cm square cardboard box to muffle the sound. The buzzer was painted white and was wired inside the box to only the two white switches. Thus the two white switches produced the sound and the three black switches did not. In this way, colour was used to decrease the memory load for this information. Before the description of the buzzer was read, children were urged to listen carefully since they would be asked to recall the description exactly.

The following description was read to each child as the experimenter demonstrated how it worked:

"This is a sound-maker. Some of these switches make a sound come from this box. If I push this white switch [pushes], you hear a sound because it's wired to the box. If I push this white switch [pushes], you hear a sound because it's wired to the box. But if I push either this black switch [pushes], this black switch [pushes], or this black switch [pushes], nothing happens. That's because they're not wired to the box."

This description was read twice and children were asked to recall the description using the same words as the experimenter. Children were instructed to manipulate the switches as they recalled the description. We used the complex "if [either-or] then" sentence (a combination of "if-then" and "either-or" constructions) because it was expected to be more complex than that typically used by young children. It was of interest whether children would "reduce" this expression to a more natural phrasing.

Scoring. Each protocol was again scored with respect to conceptual content, linguistic form, and conversions. With respect to content, the buzzer description was broken down into six causal ideas: (a) the left white switch produces the sound; (b) it produces the sound because it is wired to the box; (c) the right white switch produces the sound; (d) it produces the sound because it is wired to the box; (e) the three black switches produce no sound; and (f) they produce no sound because they are not wired. Hence, subjects could receive a score of 0–6 for content. Ideas (a), (c), and (e) are referred to in subsequent analyses as the covariance information for the buzzer task, and ideas (b), (d), and (f) are referred to as the explanation information. For linguistic form, subjects could receive a score for each of the two "if" phrases (0–2), the single "if [either-or] then" phrase (0–1), and the three "because" phrases (0–3). For "if"-to-"when" conversions, subjects could receive a score of 0–3.

Interrater reliability. After a single judge scored all protocols for the balance and buzzer tasks, a second judge independently scored a random 25% of these protocols for every grade. Agreement was 95%.

RESULTS

Results for content, form, and conversion scores for each task are analysed separately. All significant effects are interpreted using Newman–Keuls post hoc tests.

Conceptual Content. In order to determine whether there were age differences in recall of the causal ideas presented in the balance and buzzer descriptions, a 3 (grade: first, third, fifth) × 2 (task: balance, buzzer) ANOVA with repeated measures was computed. For this analysis, raw scores for each task were converted into proportions so that performances on both tasks could be assessed within the same analysis. However, since proportions reflect non-normally distributed data, arcsine transformations were computed for each subject's proportion score (Winer, 1971). The ANOVA revealed a significant main effect for grade, $F(2,75) = 17.47$, $P < 0.001$. Here, the fifth-graders (96%) recalled significantly more of the causal ideas across tasks than the third-graders (87%), who recalled significantly more than the first-graders (73%). The main effect of task and the grade × task interaction were not significant.

As a second analysis, we conducted separate one-way ANOVAs on covariance ideas and explanation ideas. With respect to recall of all six covariance ideas, the fifth- ($M = 100\%$) and third-graders ($M = 98\%$) recalled significantly more than the first-graders ($M = 88\%$), $F(2,75) = 5.06$, $P < 0.01$. The older two groups did not differ. However, this result may be largely due to a ceiling effect. Developmental differences were much more pronounced for the explanation ideas across tasks. Here, the fifth- ($M = 90\%$) and third-graders ($M = 74\%$) recalled significantly more than the first graders ($M = 49\%$), $F(2,75) = 10.07$, $P < 0.001$. The difference between the fifth- and third-graders approached but did not attain significance. Additional analyses showed that children at all ages performed better when recalling cause-general information (i.e. covariance ideas) than when recalling cause-specific information (i.e. explanations).

Presented in Table 1 is a content analysis of subject protocols for the balance and buzzer tasks. In this analysis, each subject is classified according to his or her overall *pattern* of response on each problem. For example, a child who recalled all three covariance ideas but did not recall an explanation demonstrated a "covariance-only" response pattern. For the balance task, first-graders were most likely to recall either the covariance information only (i.e. that spots 2 and 3 cause the light but spot 1 does not) or all of the covariance information and an incorrect explanation. For the fifth-graders, on the other hand, the predominant pattern was to recall the correct explanation in addition to all of the covariance information (χ^2 $(6, N = 78) = 36.47$, $P < 0.001$). Thus, with age children were not only more likely to recall an explanation, but were more likely to recall the correct explanation.

An analysis of the incorrect explanations for the balance task revealed two primary categories that we label *Spatial Contiguity* and *Weight Change*. In the spatial contiguity category, subjects misconstrued the explanation

TABLE 1
Percentages of Subjects in Recall Response Patterns for the Balance and Buzzer Tasks

Balance task	*Grade*		
Response pattern	*First*	*Third*	*Fifth*
Covariance only	35	11	0
Explanation only	0	0	0
Covariance + incorrect explanation	35	35	15
Covariance + correct explanation[a]	15	54	85
Other[b]	16	0	0

Buzzer task	*Grade*		
Response pattern	*First*	*Third*	*Fifth*
Covariance only	31	8	0
Explanation only	8	4	0
Covariance + incorrect explanation	0	0	0
Covariance + explanation[c]	31	62	85
Other	30	26	15

Notes
[a]Recall of all four causal ideas.
[b]Partial recall of covariance and explanation ideas.
[c]Recall of all 6 causal ideas.

statement about "distance from the fulcrum" to mean "closer proximity to the light apparatus". For example, one first-grader said, "Number 3 makes the light 'cause it's the end one—it's close to this [the light]", and one third-grader said, "And 1 doesn't, 'cause the weight is farther away from the light". Of the first-, third- and fifth-graders' incorrect explanations, 56%, 22% and 0% respectively were of this type. Subjects in the weight-change category argued that the ring-shaped weight actually became heavier as it moved towards spot 3; i.e. they failed to recognise the invariance of its weight. For example, one fifth-grader said, "if you put it on spot number 3, it's even farther out so that the weight changes even more and keeps it down." Of the first-, third- and fifth-graders' incorrect explanations, 22%, 67% and 100% respectively were in this category. These two categories accounted for 86% of all incorrect explanations.

A content analysis of the buzzer task revealed that whereas 31% of the first-graders recalled only the covariance information (i.e. that the white switches produce sounds but the black switches do not), only 8% of the

third-graders and none of the fifth-graders limited their descriptions to the covariance ideas. Analysis of these patterns reveal significant age trends, χ^2 $(6, N = 78) = 22.27$, $P < 0.001$. Of note is the fact that no incorrect explanations were recalled for the buzzer apparatus. The differences between tasks in recall of incorrect explanations is significant, $F(1,75) = 30.71$, $P < 0.001$.[1] The greater recall of explanations of the buzzer task could be due either to better comprehension of the causal sequences in the buzzer task or to the fact that children heard the explanation three times in the buzzer description and only once in the balance description.

As to the issue of consistency, the correlation between balance and buzzer total content scores after partialling age is $r = 0.20$, n.s. This non-significant correlation appears to be largely due to inconsistencies in recall of explanation ideas. In particular, whereas 91% of subjects showed the same pattern of recall across tasks for covariance ideas, only 49% show the same pattern of recall across tasks for explanation ideas. These findings and those regarding differential recall of incorrect explanations across tasks seem to support the distinction between cause-general and cause-specific aspects of causal knowledge.

Linguistic Form. Three separate ANOVAs were computed for the incidence of "if", "because", and "if [either-or] then" statements found in children's recalled descriptions of covariance and explanation ideas. Again children's scores were first converted to proportions and then transformed via arcsine transformations in order to include task as a repeated measures factor.

For "if", the ANOVA revealed a significant main effect for grade, $F(2,75) = 3.49$, $P < 0.05$. Here, the fifth-graders used a significantly higher proportion of "if" statements (75%) to describe the covariance information than either the third- (50%) or first-graders (56%). The first- and third-graders did not differ. The main effect of task and the grade × task interaction were not significant.

The analysis for "because" also revealed a main effect for grade, $F(2,75) = 18.14$, $P < 0.001$. Here, the fifth-graders used significantly more "because" statements for explanations (78%) than the third-graders (60%), who used significantly more than the first-graders (40%). The main effect of task and the grade × task interaction were not significant.

A third analysis considered the recall of the "if [either-or] then" con-

[1]The former analysis on total content scores revealed no main effect of task because a score of "0" was awarded to both "no explanation" and "incorrect explanation" responses. This scoring obscured intertask differences.

TABLE 2
Percentages of Subjects Using Each Type of Linguistic Construction on the Balance Task

	Grade		
Covariance	*First*	*Third*	*Fifth*
"If" exclusively	46	38	69
Action commentary exclusively	23	24	4
Mixture[a]	15	19	15
Other[b]	16	19	12
	Grade		
Explanation	*First*	*Third*	*Fifth*
"Because" exclusively	12	42	77
Action commentary[c] exclusively	8	16	8
No explanation	80	42	15

Notes:
[a]Contains a mixture of "if" and action commentaries.
[b]Reflects children who either failed to recall or who used "when" instead of "if".
[c]Also contains conversions of "because" to "so".

struction found only in the buzzer task description. Contrary to the above analyses, this ANOVA revealed no grade effect (first, 35%; third, 35%; fifth, 39%). Again, neither the main effect of task nor the grade × task interaction were significant.

Tables 2 and 3 show the types of linguistic constructions used by children to describe covariance and explanation information for the balance and buzzer tasks. As can be seen in Table 2, there was an increasing tendency with age to use "if" statements exclusively to describe the covariance ideas on the balance task accompanied by a decreasing tendency to use "action commentaries" exclusively to describe this information. Action commentaries consisted of mostly declarative sentences which did not contain "if" and which described the actions that subjects were performing. Examples are: "Put it on number 1?—nothing happens" and "This one makes the light go on." Inspection of Table 2 also suggests that the primary reason that first-graders failed to use "because" on the balance task was that they failed to recall an explanation. By the fifth grade, however, most children attempted an explanation and most used "because" exclusively.

With regard to the buzzer task, Table 3 shows that children at every grade were equally likely to use "if" exclusively to describe both active switches. It can also be seen that children at all grades were equally

TABLE 3
Percentages of Subjects Using Each Type of Linguistic Construction on the Buzzer Task

	Grade		
Covariance (active switches)	*First*	*Third*	*Fifth*
"If" exclusively	54	50	62
Action commentary exclusively	23	35	23
Mixture	12	12	15
Other	11	3	0
	Grade		
Covariance (inactive switches)	*First*	*Third*	*Fifth*
"If [either–or] then"	35	35	39
Other "if" statements	15	31	30
Action commentary exclusively	50	34	27
	Grade		
Explanation	*First*	*Third*	*Fifth*
'Because" exclusively	27	50	73
Exclusive action commentary	27	8	0
Mixture	19	31	27
No explanation	27	11	0

*un*likely to use the "if [either-or] then" phrasing to describe the three inactive switches. They were more likely to use action commentaries, or to use an "if any of these" construction, which was perfectly synonymous with "if [either-or] then". The low incidence of the latter construction may reflect the fact that it was too complex or long to process. The tendency to reduce this construction to a simpler form suggests that older children were not mindlessly mimicking constructions.

For "because", the primary finding listed in Table 3 is that only 27% of the first-graders failed to recall some part of an explanation (first column, last row) as compared to the balance task where 80% failed to explain. Relatively the same number of third- and fifth-graders, however, recalled explanation ideas for both tasks, and relatively the same number predominantly used "because".

Whereas Tables 2 and 3 illustrate the types of syntactic constructions recalled by children and shed some light on the reasons why significant age effects were found for "if" and "because" statements, these Tables do not indicate whether the grade differences for these connectives were due to

the differential use of "if" over other constructions at certain grades, or simply to the fact that older children recalled more causal ideas (as the content analysis shows). To clarify the issue, each subject's recall scores for content and form were used to compute conditional probabilities for the use of "if" and for the use of "because". These scores consisted of the probability that a subject used "if" given that he or she recalled covariance ideas, and the probability that he or she used "because" given that explanations were recalled. First, scores across tasks were collapsed into a single conditional probability. Then arcsine transformations were computed for each probability score. A one-way ANOVA on the transformed conditional probabilities for "if" revealed a significant effect for grade, $F(2,75) = 2.98$, $P < 0.05$. Here, the fifth-graders ($M = 0.75$) differed significantly from the third-graders ($M = 0.51$), though not from the first-graders ($M = 0.62$).

These analyses for conditional probabilities suggest that the significant difference in recall of "if" between the first and fifth-graders can be explained differently from the difference in recall of "if" between the third- and fifth-graders. It would appear from the analysis of conditional probabilities that the first-graders used "if" less than the fifth-graders because they were less likely to recall covariance ideas (88% vs 100%). The third-graders however, differed from the fifth-graders in the earlier analysis not because they recalled fewer covariance ideas (98% vs 100%), but because they were more likely to use other syntactic constructions (e.g. action commentaries) besides "if" to describe the ideas they recalled.

A one-way ANOVA for the conditional probabilities for "because" also revealed a significant grade effect, $F(2,67) = 6.31$, $P < 0.003$. In this case, the fifth-graders ($M = 0.88$) differed significantly from the third-graders ($M = 0.73$), who differed significantly from the first-graders ($M = 0.54$). Since the same age differences were found for conditional probabilities as for raw scores, we conclude that there is a genuine increase with age in memory for "because" over other linguistic constructions. Even when the first-graders recalled explanations, they used "because" only 54% of the time.

An important finding concerning linguistic form, however, was that no child at any grade ever made a reversed-clause error for "if" or "because". That is, children never said things like "The weight is far out because the light is on."

Conversions. A final analysis for the balance and buzzer tasks considered the frequency of "if"-to-"when" conversions during recall. The analysis for the balance task revealed no significant effect for grade, $F(2,75) = 0.30$. The mean number of conversions were 0.50, 0.58, and

0.31 (out of a possible 3.0) for the first-, third-, and fifth-graders, respectively. The corresponding percentages of children who made at least one conversion were 31%, 23%, and 12% respectively. For the buzzer task, there were again no developmental differences, $F(2,75) = 1.03$. No first-grader, two third-graders, and one fifth-grader made an "if"-to-"when" conversion on the buzzer task. These findings suggest that children generally do not find "if" and "when" to be interchangeable in this context.

DISCUSSION

Experiment I employed a recall methodology to examine developmental differences in (a) memory for the covariance vs explanation information presented in each task description, (b) the incidence of "if" and "because" in subjects' recall protocols, and (c) the incidence of reversed-clause errors. Results showed that different developmental trends emerged for recall of the covariance of causes and effects than for recall of explanations of causal mechanisms. In particular, the differences between the third- and fifth-graders in recall was much more pronounced for explanation information than for covariance information. These findings are consistent with, but by no means unequivocally supportive of, our differential knowledge hypothesis. That is, the covariance information can be considered cause-general, whereas the explanation information can be considered cause-specific. We consider the interpretation of these results in the general discussion.

Second, whereas developmental differences emerged with respect to use of the connectives "if" and "because" in subjects' recalled versions of the descriptions, these differences need to be interpreted in the light of subjects' recall of covariance and explanation ideas. When one controls for recall of covariance ideas using conditional probabilities, the difference between first- and fifth-graders in absolute frequency of "if" used is no longer significant. However, this was not the case for "because". Even after one controls for children's recall of explanation ideas, the fifth-graders were still significantly more likely to use "because" within explanations than were the first- and third-graders. These findings suggest that children may have a greater variety of expressions for explanations than for covariance relations.

Finally, whereas Emerson (1979, 1980) reported a high incidence of reversed-clause errors in even her third- and fifth-graders using a judgment task, no children in the present study made such an error. Hence, the recall methodology proved to be more sensitive to children's competence than a task which required metalinguistic judgements.

EXPERIMENT II

The primary goal of experiment II was to employ a recall methodology to assess children's knowledge of the pragmatic concepts conveyed by "if" and "when". It was expected that such knowledge would be more probably revealed using a recall methodology, since metalinguistic judgements regarding presuppositions were not required.

Method

Subjects. The same first-, third-, and fifth-graders who participated in experiment I participated in experiment II.

Tasks and Procedure

Sentence Recall Task. Following the balance and buzzer tasks, children were read six sentences, three of which were constructed with "if", and three of which were constructed with "when". Half of the sentences had semantic content which referred to unusual or unlikely events, and half referred to events that are more typical and familiar to school children. Subjects were randomly assigned to one of two conditions formed by pairings of connectives and content. In one condition, the unusual semantic content was preceded by "if' and the typical semantic content was preceded by "when". We refer to this condition as the "if-unusual" condition. Since "if" is usually used to convey uncertainty to the listener (Scholnick & Wing, 1982), we attempted to maximise the possibility that children would recognise uncertainty by having "if" describe unusual events. Certainty and uncertainty can be thought of as a continuum ranging from (a) "impossible" at one extreme, moving to (b) "very unlikely", to (c) "somewhat unlikely", to (d) "no knowledge", to (e) "somewhat likely", to (f) "very likely", and finally to (g) "certain" at the other extreme. In English, "if' with the subjunctive mood is used for events that the speaker knows are false or impossible (e.g. "If Carter had been re-elected . . ."), and "because", "when", and "so" with the indicative mood for events the speaker believes to be certain (e.g. "When I'm older . . ."). A speaker uses "if" with the indicative mood for all other cases on the continuum (see Scholnick & Wing, 1982, for a more detailed discussion). We employed the indicative mood specifically for all sentences so that the presence of "if" vs "when" in a sentence could not be predicted by sentence mood. To consider how "if" conveys relatively more uncertainty than certainty and "when" conveys relatively more certainty than uncertainty, compare "If it rains today, the game will be cancelled" and "When it rains today, the game will be cancelled."

Half of the subjects at each grade were placed in the "if-unusual" condition (where "if" was paired with unusual content) and heard the following sentences:

(a) If Michael Jackson walks in, we'll be surprised.
(b) If I see a dog flying, I'll get my eyes checked.
(c) If summer starts tomorrow, I'll be happy.
(d) When I'm tired, I like to sleep.
(e) When it's Saturday, I like to watch TV.
(f) When it's night-time, I turn on the lights.

Since the pairing of content and connectives in this condition was consistent with normal usage, very few conversion errors were expected. Hence, it was expected that subjects would recall the unusual content of sentences (a)–(c) using "if" and the more typical content of (d)–(f) using "when".

To test the subjects' knowledge of the association of "if" with the less likely content and "when" with the more likely content, half of the children in each grade were randomly placed in the "when-unusual" condition. Here, "when" was paired with the same unusual content in (a)–(c) above (e.g. "When Michael Jackson walks in, we'll be surprised") and "if" was paired with the same normal content of (d)–(f) above (e.g. "If I'm tired, I like to sleep"). If children expect "if" to signal unusual content and "when" to signal usual content they should produce many conversion errors in the "when-unusual" condition and few conversion errors in the "if-unusual" condition. It should be noted that the content of sentence (c) was unusual in that subjects were tested in midwinter.

Subjects were asked to recall the sentences after the experimenter read aloud all six in a random order. Subjects were told that they would hear six sentences read in a row and were to listen carefully since they would be asked to remember all six sentences "word for word after I read all six of them." A cued recall procedure was used where a key word or phrase in each sentence was used as the cue. After all sentences had been read, the experimenter would say, for example, "Tell me the one about Michael Jackson." The other cues were "flying dog", "summer", "being tired", "Saturday", and "night-time" for sentences (b) through (f) respectively.

Scoring. Each subject's sentence recall was scored for content, form, and conversions. With respect to content, subjects received one point for a given sentence only if they correctly recalled the ideas found in both clauses of that sentence. Thus, recalling only the first clause which contained the cue, or recalling separate ideas found in different sentences received a score of zero. Close paraphrases (e.g. "If Michael Jackson steps

in . . .") were permitted for the content scoring. Subjects could receive a score of 0–6 for content. For linguistic form, subjects received one point for each sentence in which they recalled "if" and "when" correctly (as determined by their condition). For example, subjects in the "if-unusual" condition received 6 points if they recalled the unusual content with "if" and typical content with "when". Subjects who omitted the connective completely or who converted one connective into another received a score of zero. For conversions, subjects received one point for each "if"-to-"when" or "when"-to-"if" conversion as determined by their condition. An "if-to-when" conversion occurred if a subject heard "if" paired with a given content but recalled this content using "when", and a "when-to-if" conversion occurred if a subject heard "when" paired with a given content but recalled this content using an "if" sentence. Subjects could receive a score of 0–6 for conversions.

RESULTS

Analyses were performed for content, linguistic form, and conversion scores considered separately. All post hoc analyses used Newman–Keuls tests.

Content. In order to assess age and condition differences in recall of sentence content, a 3 (grade: first, third, fifth) × 2 (condition: "if-unusual", "when-unusual") ANOVA was computed. This analysis revealed a main effect for grade, $F(2,72) = 4.25$, $P < 0.02$. The main effect of condition and the grade × condition interaction were not significant. Analysis of the grade effect for content revealed that the fifth-graders recalled significantly more of the clause ideas than the first-graders. No other mean differences were significant (*M*s = first, 68%; third, 73%; fifth, 83%).

Linguistic Form. In order to assess age and condition differences in the ability to recall the exact connectives used by the experimenter, a 3 (grade: first, third, fifth) × 2 (condition: "if-unusual", "when-unusual") ANOVA was computed for form scores. This analysis revealed significant main effects for grade, $F(2,72) = 5.65$, $P < 0.006$, and condition, $F(1,72) = 44.72$, $P < 0.001$. For the grade effect, the fifth-graders ($M = 62\%$) were more likely to pair connectives with content during recall in the same way that they heard it than either the third- ($M = 46\%$) or first-graders ($M = 46\%$). For the main effect of condition, the children in the "if-unusual" condition ($M = 66\%$) were more likely to remember the way the content and connectives were paired than children in the "when-unusual" condition ($M = 36\%$). Thus, it appears that the unconventional

pairings of "when" with the unusual content and "if" with the more typical content produced a decrement in performance for linguistic form scores, and that the more conventional pairings of "if" with the unusual content and "when" with the typical content facilitated recall of the precise wording. The lack of the grade × condition interaction suggests that the decrement in performance was present at every grade level.

Conversions. A 3 (grade) × 2 (condition) ANOVA assessed grade and condition differences in the incidence of "if"-to-"when" and "when"-to-"if" conversions. This analysis revealed no main effect for grade, but a significant main effect of condition, $F(1,72) = 81.58$, $P < 0.001$, and a grade × condition interaction, $F(2,72) = 8.88$, $P < 0.001$. Whereas the third- and fifth-graders were significantly more likely to make conversions in the "when-unusual" condition (*M*s = third, 50%; fifth, 50%) than in the "if-unusual" condition (*M*s = third, 6%; fifth, 4%), the first-graders made a fairly equal number of conversions in both conditions (*M*s = "when-unusual", 32%; "if-unusual", 21%). Thus, children in the older two grades were nearly ten times more likely to convert the wording when the connectives and content were paired in an unconventional way than when paired in the conventional way. This shows that they make a principled distinction between "if" and "when". Additional post hoc tests revealed that the third- and fifth-graders were significantly less likely than the first-graders to convert in the "if-unusual" condition, but were significantly *more* likely than the first-graders to convert (i.e. "err") in the "when-unusual" condition.

Table 4 lists the conditional probabilities that subjects recalled the connectives in the same way as the experimenter ("preserved pairing"), converted ("reversed pairing"), or used no connective, given that the correct content was recalled. ANOVAs on "preserved pairing" and "reversed pairing" conditional probabilities revealed essentially the same findings as the former analyses of linguistic form and conversion scores, since these scores were assigned only if the subject recalled the content. An important difference, however, is that now the first-graders showed a significantly higher probability of converting in the "when-unusual" condition than they did in the "if-unusual" condition. As to the "no connective" category, an ANOVA revealed a main effect of grade, $F(2,75) = 4.39$, $P < 0.02$, but no main effect of condition nor a grade × condition interaction. Here, the first-graders were significantly more likely to recall the content without a connective (21% of the time) than either the third- (8%) or fifth-graders (7%).

In sum, then, the first-graders were (a) less likely to recall the connective-content pairings than fifth-graders, (b) more likely to convert in the "if-unusual" condition and less likely to convert in the "when-unusual"

TABLE 4
Sentence Recall Task: Conditional Probabilities

	Condition					
	If-unusual			*When-unusual*		
	1st	*3rd*	*5th*	*1st*	*3rd*	*5th*
Preserved pairing	0.60	0.76	0.84	0.34	0.35	0.46
Reversed pairing	0.24	0.16	0.04	0.40	0.57	0.53
No conjunction	0.16	0.08	0.12	0.26	0.08	0.01

Note: "Preserved pairing" means that connectives and content were paired during recall in the same way as they were presented. These are conditional probabilities for linguistic form scores. "Reversed pairing" indicates conditional probabilities for conversions.

condition than the third- and fifth-graders, and (c) more likely to recall the content without a connective. These findings suggest a lesser ability to distinguish among "if" and "when" in the first-graders. However, it is a matter of degree as opposed to an absolute lack of ability. The atypical pairing of "when" with the unusual content tended to increase the probability that first-graders would forget the content. When they did recall the content, they showed a greater tendency to convert.

GENERAL DISCUSSION

The primary purpose of the present study was to determine how children's knowledge of causal relations and knowledge of linguistic aspects of causal expressions might affect their memory for causal descriptions. We employed a recall methodology wherein we hypothesised that children may have differing levels of knowledge of causal sequences and causal expressions (as suggested by prior research) and that this would be reflected in which aspects of the descriptions they remembered. Additionally we were interested in the linguistic forms children used in their recall to assess the frequency of "if" and "because" as well as of ill-formed constructions such as reversed-clause errors.

There were three primary findings in the present study. The first was that there were different developmental trends in recall of covariance ideas than in recall of explanation ideas on the balance and buzzer tasks. In particular, whereas recall of covariance ideas was quite high at all grades, recall of explanation ideas showed more marked developmental differences. Overall, whereas the probability that any one of our subjects would recall all the covariance ideas is 0.95, the probability that any one subject would recall the correct explanations is 0.69. What do these differences

mean? A first proposal is that older children simply have greater short-term memory capacity at their disposal than younger children. However, short memory capacity is typically thought to be indifferent to content. Hence, if the capacity of a given first-grader is say, 3 units, technically any three ideas could have been recalled. The fact that in the majority of cases where children omitted ideas they omitted explanation ideas (see Table 1) argues against a simple increased capacity proposal.

A second proposal is that these results are an artifact of the structure of the descriptions; that is, there is something about the structure of the descriptions which produced the differential recall of covariance and explanation ideas. However, the balance and buzzer descriptions were structurally distinct. The balance description had the explanation placed at the end which should have maximised recall due to recency. The buzzer description had portions of the explanation interspersed throughout and certain portions repeated twice. The fact that the same differential recall patterns occurred for both tasks argues against a simple artifact account.

The third proposal, which we favour, specifies that the differential patterns of recall for covariance and explanation ideas reflect differential knowledge of specific aspects of causal relations. In particular we based our differential recall predictions on prior research which showed that very young children evidence accurate knowledge of cause-general principles of causal attribution such as *covariance* or *temporal order* (Bullock et al., 1982), but often have difficulty when asked to explain causal sequences (Berzonsky, 1971; Byrnes & Beilin, 1987 (Note 1); Siegler, 1976). Explanations constitute cause-specific knowledge which necessarily derives from experience with given sequences and from attempts at constructing a model specifying a mechanism linking cause and effect. Our findings are analogous to those of Siegler (1976), who found that children who fail to consider distance in their predictions about a balance also fail to consider distance when asked to recall the locations of weights on a given balance. However, we need to point out that our data only *indirectly* support this third proposal, since we did not independently assess knowledge and relate knowledge scores to recall scores. Much stronger support for the differential knowledge hypothesis would come from a study which correlated knowledge with recall using subjects of a single age-group who differed in knowledge of the causal relations presented. Until such a study is conducted, we can only say that the present data are consistent with our hypothesis.

However, a particular version of the differential knowledge proposal, constructivist theory, achieved somewhat more direct support from our data regarding incorrect explanations. In this view, memory is not simply a copy of experience. Rather, representations in memory are constructions and the knowledge they represent at any particular time constrains what is

assimilated (Piaget & Inhelder, 1973). Constructivists further argue that present knowledge can lead to distortions or deletions based on a tendency to schematise or summarise information in line with a current conceptualisation (e.g., Bartlett, 1932; Piaget & Inhelder, 1973). Furthermore, knowledge in memory forms the basis of inferences which fill in gaps in the material presented (Paris, 1978). In the present study some children distorted the explanation regarding distance on the balance task to one of spatial proximity, and others favoured a conceptualisation where objects change their weight as they move farther from the fulcrum. These distortions were obviously not simple imitations of what was said. However, because we did not independently assess knowledge, incorrect explanations during recall cannot be unequivocally ascribed to deficiencies in knowledge.

The second primary finding concerns children's knowledge of the linguistic form of causal expressions. As predicted, there were developmental increases in the frequency of "if" and "because" statements to describe causal sequences. This increase was not simply a function of the amount recalled. The analysis of conditional probabilities that considered the probability that "if" or "because" was used, given that the content was recalled, revealed continued developmental differences for "because" (though inconsistent developmental differences for "if").

The third major result was found for children's implicit pragmatic marking of "if" (Bates, 1976) for unusual or unexpected events. The results for the sentence recall task showed that the third- and fifth-graders systematically paired "if" with unusual content and "when" with the more typical content regardless of how they heard these connectives and contents paired by the experimenter. Even first-graders showed, however, better recall in the "if-unusual" condition than in the "when-unusual" condition and showed a higher probability of converting in the latter condition when one controls for recall of content. These findings strongly suggest that "if" is marked for unusual or unexpected events and "when" is marked for ordinary or expected events, at least implicitly, by the first grade. Consistent with this claim is the fact that very few children converted "if" to "when" in their recall of the balance and buzzer descriptions. The findings for the sentence recall task can be said to provide indirect support for a constructivist perspective in that the recall data suggest that children cognitively altered linguistic input in line with linguistic schemata in the same way that Bransford & Franks' (1972) subjects did.

It is also important to note that no child at any grade made a reversed-clause error. This finding suggests that children at every grade have the implicit knowledge of the temporal order constraint on "if" and "because" expressions. This finding replicates those of studies which examined chil-

dren's spontaneous "if" and "because" expressions in natural contexts (e.g., Hood & Bloom, 1979). However, since the causal sequences were observable and non-reversible, our tasks could be thought of as providing semantic and contextual supports to the syntactic constraint of temporal order (Emerson, 1980). Even with such supports, however, the extremely high level of canonical order utterances argues against the claim that children below the age of 8 years have no knowledge of temporal order constraints.

These data then are very consistent with research both on children's knowledge of causal relations and their facility with causal expressions. The recall methodology employed appears to be a successful approach, and could easily be used to conduct a more fine-grained analysis than presented here. In particular, analogous to the sentence recall task, it would be interesting to assess the effects of having children recall descriptions of apparatuses which were conceptually erroneous or descriptions which contained statements violating the temporal order constraint, to see if children would recall these erroneous descriptions in the conceptually and linguistically correct fashion. Again, however, one would have to measure both knowledge and recall to provide strong, direct support for either the differential knowledge or constructivist perspectives.

A final point is that the present age differences cannot be accounted for simply by saying that older children are better able to follow instructions. Older children actually perform *worse* during recall than younger children when input is counter to their schema. These patterns and the fact that our findings are entirely consistent with prior research suggest the utility of a knowledge-based approach to memory and cognition.

Manuscript accepted 4 April 1989

REFERENCES

Bartlett, F. C. (1932). *Remembering*. Cambridge: Cambridge University Press.

Bates, E. (1976). *Language in context*. New York: Academic Press.

Berzonski, M. D. (1971). The role of familiarity in children's explanations of physical causality. *Child Development, 42*, 705–715.

Bloom, L. & Capatides, J. B. (1987). Sources of meaning in the acquisition of complex syntax: The sample case of causality. *Journal of Experimental Child Psychology, 43*, 112–128.

Bransford, J. D. & Franks, J. J. (1972). The abstraction of linguistic ideas. *Cognition, 2*, 211–249.

Brown, A. L., Bransford, J. D., Ferrara, R. A., & Campione, J. C. (1983). Learning, remembering and understanding. In P. Mussen (Ed.), *Handbook of child psychology* (Vol. III). New York: John Wiley & Sons. (Pp. 77–166.)

Brown, R. (1973). *A first language*. Cambridge, Mass.: Harvard University Press.

Bullock, M., Baillargeon, R., & Gelman, R. (1982). The development of causal reasoning. In W. Friedman (Ed.), *The developmental psychology of time*. New York: Academic Press. (Pp. 209–254.)

Chi, M. T. H. (1978). Knowledge structures and memory development. In R. S. Siegler (Ed.), *Children's thinking: What develops?* Hillsdale, N.J.: Lawrence Erlbaum Associates Inc. (Pp. 73–176.)

Emerson, H. F. (1979). Children's comprehension of "because" in reversible and non-reversible sentences. *Journal of Child Language*, *6*, 279–300.

Emerson, H. F. (1980). Children's judgments of correct and reversed sentences with "if". *Journal of Child Language*, *7*, 137–155.

Flavell, J. H. (1985). *Cognitive development*. Englewood Cliffs, N.J.: Prentice-Hall.

French, L. A. & Nelson, K. (1985). *Young children's knowledge of relational terms: Some ifs, ors and buts*. New York: Springer-Verlag.

Hood, L. & Bloom, L. (1979). What, when, and how about why: A longitudinal study of early expressions of causality. *Monographs for the Society for Research in Child Development*, *44* (6), (Serial No. 181).

Lindberg, M. A. (1980). Is knowledge base development a necessary and sufficient condition for memory development? *Journal of Experimental Child Psychology*, *30*, 401–410.

Paris, S. G. (1978). The development of inference and transformation as memory operations. In P. A. Ornstein (Ed.), *Memory development in children*. Hillsdale, N.J.: Lawrence Erlbaum Associates Inc. (Pp. 129–156.)

Piaget, J. (1974). *Understanding causality*. New York: Basic Books.

Piaget, J. & Inhelder, B. (1973). *Memory and intelligence*. New York: Basic Books.

Scholnick, E. K. & Wing, C. S. (1982). The pragmatics of subordinating conjunctions: A second look. *Journal of Child Language*, *9*, 461–479.

Scholnick, E. K. & Wing, C. S. (1983). Evaluating presuppositions and propositions. *Journal of Child Language*, *10*, 639–660.

Shultz, T. R. (1982). Rules of causal attribution. *Monographs of the Society for Research in Child Development*, *47* (1), (Serial No. 194).

Shultz, T. R. & Mendelson, R. (1975). The use of covariation as a principal of causal analysis. *Child Development*, *46*, 394–399.

Siegler, R. S. (1976). Three aspects of cognitive development. *Cognitive Psychology*, *8*, 481–520.

Siegler, R. S. (1980). When do children learn? The relationship between existing knowledge and learning. *Educational Psychologist*, *15*, 135–150.

Siegler, R. S. (1986). *Children's thinking*. Englewood Cliffs, N.J.: Prentice-Hall.

Slobin, D. I. (1979). *Psycholinguistics* (2nd edn.). Glenville, Ill.: Scott, Foresman & Co.

Winer, B. J. (1971). *Statistical principles of experimental design* (2nd edn). New York: McGraw Hill.

REFERENCE NOTE

1. Byrnes, J. P. & Beilin, H. (1987). *The relation between causal and logical reasoning in children*. Paper presented at the Biennial Meeting of the Society for Research in Child Development, Baltimore, Md.