

REPLY

How to Detect Reasoning-Remembering Dependence
(And How Not To)

MICHAEL CHAPMAN†

University of British Columbia

AND

ULMAN LINDENBERGER

Max Planck Institute for Human Development and Education

In opposition to Brainerd and Reyna (1992), it is argued that the absence of an association between children's reasoning and memory for premises in a number of studies does not reflect any underlying independence between reasoning and remembering. Instead, those nonsignificant results have occurred either (a) because memory was tested for premises other than those actually used in reasoning or (b) because the proportion of children reasoning from the premises actually tested was insufficient to reject the null hypothesis of independence. When these factors are controlled, a dependence between reasoning and remembering is indeed found, as described by Chapman and Lindenberger (1992). Brainerd and Reyna's claim that the data reported in that study show reasoning-remembering independence when reanalyzed is found to be based on confusion between *stochastic* and *deterministic* dependence and on faulty statistical arguments. © 1992 Academic Press, Inc.

In our view, Brainerd and Reyna's (1992) comments on our article, "Transitivity Judgments, Memory for Premises, and Models of Children's Reasoning" (Chapman & Lindenberger, 1992) are characterized by three confusions: (a) a confusion between issues of memory and issues of logical form, (b) a confusion between two senses of the term "dependence," stochastic and deterministic, and (c) a confusion between our operational constructive model of children's reasoning and what they call the classic, "logicist" view—that transitivity problems in general must be solved by deduction from pairwise comparisons between adjacent terms. Once these confusions are cleared up, Brainerd and Reyna's specific criticisms of our article can more easily be answered.

MEMORY FOR WHICH PREMISES?

Imagine an elementary school teacher who believed that achievement

Address all reprint requests to Ulman Lindenberger, Max Planck Institute for Human Development and Education, Lentzeallee 94, 1000 Berlin 33, Germany.

† Deceased.

motivation could be enhanced by making performance more salient and who changed the seating arrangement in the classroom after every examination so that children were ordered from left to right in order of increasing grade. Then by observing the new seating arrangement, the children could infer their relative standing on the exam. The form of such an inference would be as follows:

- A1. Grades increase from left to right.
- A2. Gwen is sitting to the right of Daryl.
- A3. So Gwen must have gotten a higher grade than Daryl did.

Suppose further that a new child enters the class. Ignorant of the special seating assignments, the new child is unable to infer relative standing from the spatial arrangement. By comparing grades with two classmates, however, the new child reasons as follows:

- B1. Gwen got a higher grade than I did.
- B2. I got a higher grade than Daryl.
- B3. So Gwen must have gotten a higher grade than Daryl did.

Although both forms of inference lead to the same conclusion ($A3 = B3$), those conclusions derive from different sets of premises ($A1-A2$ and $B1-B2$, respectively).

Finally, suppose that a psychologist comes into the classroom, tests children's memory for B1 and B2 as well as their ability to infer the relative standing of Gwen and Daryl, finds no statistically significant relation between the two variables, and concludes in a research article that, *in general*, reasoning is independent of memory for premises. A reviewer of that article might object that the author has overgeneralized from the results: For one thing, there are well-known statistical problems associated with generalizing from nonsignificant results. For another, the independence of judgments from memory for B1-B2 provides no support for the conclusion that reasoning is independent of memory for premises *in general* if the children inferred their judgments from different premises entirely (say, A1-A2).

In our original article on reasoning-remembering dependence (Chapman & Lindemberger, 1992), we assumed the role of such a reviewer with respect to Brainerd and Kingma's (1984, 1985) and Brainerd and Reyna's (1990) arguments for the independence of reasoning and memory for premises. For simplicity, we focused on transitive reasoning, although our critique applies to other tasks as well. In brief, Brainerd and Kingma (1984) reported that children's memory for comparisons between adjacent elements in a transitive series (e.g., $A < B < C$) was stochastically independent of the inference $A < C$ in a number of studies and concluded: (a) that reasoning is independent of memory for premises and (b) that, instead of reasoning from comparisons between adjacent terms, children solve transitive reasoning problems from gist-like "fuzzy-traces" having the form, "Things get bigger toward the right."

We argued instead that one must distinguish between two different forms of reasoning. First, the propositional content of Brainerd and Kingma's (1984) fuzzy-trace model of transitive reasoning can be reconstructed as follows:

F1. Things get bigger toward the right.

F2. Stick A is on the right of stick C.

F3. So A is bigger than C.

Preferring a formally descriptive label, we called this form of reasoning "functional reasoning," referring to the functional relation between size and position in F1. Strictly speaking, F1–F3 is a specific *instance* of functional reasoning. As a generic *type*, "functional reasoning" includes any inference based on a functional relation between two variables.

The second form of reasoning to consider is a reconstruction of "transitive reasoning" as ordinarily understood:

T1. A is bigger than B.

T2. B is bigger than C.

T3. So A is bigger than C.

We called this form of reasoning "operational," because in our operational–constructive model T3 is derived from T1–T2 through a composition of the operations of comparison corresponding to T1 and T2, respectively.

Although the conclusions F3 and T3 are the same, they are inferred from a different set of premises in each case. In an earlier study (Chapman & Lindenberger, 1988), we found that children justified their judgments according to the schema F1–F3 when the lengths of the comparison objects were correlated with right–left spatial position and according to T1–T3 otherwise. With respect to issues of reasoning and memory, we argued that the reasoning–remembering independence reported by Brainerd and Kingma (1984) could have resulted from the fact that the children studied inferred their judgments from one set of premises (e.g., F1–F2) but that memory was tested for another set (T1–T2). If Brainerd and Kingma had interpreted their findings merely as showing that children do not always reason from T1 and T2, then one could only agree; but if one wishes to test the independence of reasoning and memory for premises in general, then one should test memory for the premises actually used in solving the particular task. If children reason according to F1–F3, then the relevant premises are F1 and F2; T1 and T2 are relevant only if children actually reason according to the schema T1–T3. In our view, Brainerd and Kingma (and later, Brainerd and Reyna) have confused the issue by using the term "premises" to refer solely to T1 and T2, rather than to whatever premises children actually use in a particular case.

In our later article (Chapman & Lindenberger, 1992), we tested the relation between children's memory for those premises and the inference T3 under the conditions theoretically favoring the schema T1–T3 (i.e.,

absence of a correlation between length and spatial or other perceptual cues). We also conducted a power analysis to determine what proportion of children in the sample would have to reason according to T1–T2 in order to reject the null hypothesis of stochastic independence under various combinations of memory and reasoning performance. In brief, we found that, in the only test powerful enough to reject the null hypothesis according to our model, *the null hypothesis of independence was indeed rejected*; transitivity judgments were found to be stochastically dependent on memory for premises (Chapman & Lindenberger, 1992, Table 3, three-term length task). The importance of this point should not be underestimated. In our article, we perhaps gave it insufficient emphasis because we were intent on making another point: that a substantial proportion of children might reason from individual premise comparisons without that fact being reflected in a statistically significant stochastic dependence between judgments and memory for premises (Chapman & Lindenberger, 1992, Tables 3 and 7).

In their comments on our article, Brainerd and Reyna (1992) tried to minimize our finding of reasoning–remembering dependence by arguing that we inflated the Type I error rate beyond .05 in performing a test of independence for each of our five transitivity tasks (Chapman & Lindenberger, 1992, Table 3). In fact, our one critical test (the three-term length task) was significant at the .005 level, thereby meeting a Bonferroni criterion of .01 ($= .05/5$) sufficient to guarantee an overall α level of .05. (This point argued in more detail later.)

In summary, our finding of a statistically significant dependence between transitivity judgments and memory for premises in the only test with sufficient power to detect such a dependence cannot be wished away.

STOCHASTIC VERSUS DETERMINISTIC “DEPENDENCE”

Much of Brainerd and Reyna’s (1992) article is devoted to an argument that our division of children into those using an “operational” form of reasoning (such as schema T1–T3 above) in which judgments are dependent on premise memory and those children using some other form of reasoning in which no such dependence exists is irrelevant to the question of stochastic dependency. Their main point is that judgments are not in fact stochastically dependent on memory for premises in our operational subgroup because there is no variation in either judgments or memory.

That argument confounds two distinct senses of the term “dependence”: *stochastic* or probabilistic dependence and *deterministic* dependence. By focusing only on stochastic dependence, they miss the central point of our article. The main question addressed in that article was under what conditions the theoretical deterministic dependence between reasoning and memory for premises would be translated into an observable

stochastic dependence between judgments and memory probe performance. The deterministic character of the theoretical dependence between reasoning and memory follows from the central assumptions (a) that children can reason according to the schema T1-T3 only if the premises T1 and T2 are remembered correctly and (b) that their judgments will always be correct when they do follow that schema. The first assumption is merely an explicit statement of reasoning-remembering dependence and the second reflects the relation of necessity by which T3 follows from T1 to T2. In our earlier articles, we argued that this necessity is synthetic rather than analytic, resulting from the composition of operations corresponding to T1 and T2 rather than from the composition of the meanings of those propositions. As a statement of the assumptions embodied in the hypothesis of reasoning-remembering dependence, we believe these assumptions to be unexceptionable, and Brainerd and Reyna have not questioned them as such.

Note, however, what those assumptions imply: that *for children reasoning in this way*, both judgments and memory for premises will always be correct. There will be no variation in these variables, and therefore the dependence between reasoning and remembering in this subgroup will be deterministic and not stochastic. Accordingly, one cannot test the deterministic relation between variables in the operational subgroup with statistics designed to test stochastic dependence. However, one can test whether the distribution predicted for the operational subgroup by virtue of the presumed deterministic relation between memory and reasoning actually occurs, and our prediction analyses indicated that it did.

Thus, Brainerd and Reyna's claim that we have not demonstrated stochastic dependence in our operational subgroup is true, but irrelevant to our argument. The argument is that stochastic dependence can result in the *total* population through a mixture of subpopulations in which the predicted deterministic dependence is present and absent, respectively. Brainerd and Reyna refer to Simpson's paradox—the fact that collapsing two contingency tables may result in a different relation between variables than that shown in either of the original subtables—but they fail to see its relevance for our distributional model. A particularly striking form of the paradox is the case in which contingency or stochastic dependence occurs in a mixed sample, *but not in the subsamples of which the total sample is composed*. Such a mixture is precisely what we have proposed in the case of reasoning-remembering dependence. The main point of our distributional model was to explore the conditions (proportions of operational reasoners, levels of correct judgment and memory) under which the deterministic dependence of reasoning on memory in the operational subgroup would result in an observable stochastic dependence between judgments and memory in the total sample.

Because Brainerd and Reyna missed this point, their argument that our

distributional model presupposes high rates of measurement error is fallacious. The linchpin of that argument was the claim that if measurement is perfect, then reasoning and remembering are stochastically independent in both operational and nonoperational subgroups. In making this claim, Brainerd and Reyna have forgotten Simpson's paradox, in particular, the possibility that the two variables could be stochastically related in the *combined* sample, even if they are not so related in the subgroups. The fact that such results can occur is illustrated in the data for our three-term length task; no stochastic relation between reasoning and remembering exists in the operational and nonoperational subgroups taken separately (Chapman & Lindemberger, 1992, Table 4), but such a relation is indeed found in the combined sample (Table 3). Far from presupposing measurement error, our power analysis was computed under the assumption of *perfect* separation of operational and nonoperational subgroups. As we stated in our article, the existence of measurement error in identifying those groups is likely to make the detection of reasoning-remembering dependence in the data more difficult than is implied in that analysis.

Third variables. A related criticism raised by Brainerd and Reyna (1992) is that the relation between judgments and memory found in the operational subgroup is a third-variable artifact resulting from the use of verbal justifications for identifying that group. As potential candidates for such a third variable, they mention verbal fluency and motivation, but other candidates might be imagined as well.

We agree that because the operational subgroup represents a higher developmental level than the nonoperational group, that group is indeed likely to be higher on some other abilities as well. Whether those other abilities can *account* for the observed relation between judgments and memory is another question. The main reason why we believe such a third-variable explanation is unlikely is because our distributional model implies a more specific and more stringent relation between judgments and memory than would be expected according to most third-variable explanations. The deterministic relation between judgments and memory hypothesized to exist in the operational subgroup implies that both judgments and memory for premises should be *perfect*, except for the possibility of measurement error. In fact, zero errors were obtained on all five transitivity tasks (Chapman & Lindemberger, 1992, Tables 4 and 5). According to a third-variable model, one would expect a spurious *association* between the two main variables of interest, but not necessarily perfect performance on both of them. (—unless one has an additional argument to explain why the third variable should result in error-free judgments and memory; but it is difficult to imagine any plausible argument of this kind for variables such as verbal fluency, motivation, etc.)

The perfect judgment and memory performance observed in the operational subgroup also makes it impossible in principle to demonstrate that any proposed third-variable model is correct. If an association between A and B is the spurious result of a third variable C, then the association between A and B should disappear when C is controlled for, that is, when the $A \times B$ contingency table is partitioned by C. However, because judgments and memory are perfect in the operational subgroup, partitioning that subgroup according to any proposed third variable will still result in perfect performance in both groups. Therefore, we conclude that an appeal to third-variable models to explain this finding has the status of an ad hoc hypothesis that could be raised against almost any association found between two variables in developmental research.

The surface plausibility of the third-variable argument depends in part on the claim that the criterion used to partition the sample into operational and nonoperational was a highly stringent one. Thus, Brainerd and Reyna state that "on the average" 97 of 120 children could not meet the criterion. But that figure was obtained by averaging over our three-, four-, and five-term length tasks and the three- and four-term weight tasks (Chapman & Lindenberger, 1992, Table 4). We believe that averaging over these heterogeneous cases makes little sense in this context. Most theorists would agree that increasing the number of terms in a transitivity task increases its memory demand over and above whatever minimal abilities are required to find an operational explanation as such. A more revealing figure is that 47.5% of the sample was able to give an operational justification on at least one of the tasks. By averaging over tasks with such added memory demand, one inflates the Type II error and underestimates children's competence in the manner against which Brainerd (e.g., 1973) has inveighed so eloquently over the years.

Circular reasoning? A third model-related criticism mentioned by Brainerd and Reyna (1992) is that our power analysis of the conditions under which reasoning—remembering dependence might be detected involves "circular reasoning" because it assumes the truth of our distributional model. But the same charge could be leveled against any instance of hypothesis testing whatsoever, because all hypothesis testing involves a determination of what outcomes are likely to occur *if the hypothesis were true*. Only then can one test to see whether those outcomes actually do occur. Our power analysis merely illustrates an elementary statistical principle: that nonsignificant results (e.g., the absence of significant associations between reasoning and remembering) do not imply the truth of the null hypothesis, because nonsignificant results can occur for various reasons besides the absence of an effect.

In our previous article, we demonstrated that, for certain levels of memory and judgment, the null hypothesis of reasoning—remembering

independence would not be rejected even with a relatively large proportion of operational reasoners in the sample. When the proportion of operational thinkers on a given transitivity task exceeded the criterial proportion theoretically needed to detect reasoning–remembering dependence, such dependence indeed was found. Otherwise, the results were nonsignificant, even with an appreciable proportion of operational thinkers (up to 28%) for the task in question (Chapman & Lindemberger, 1992, Table 7). The import of these results is that, unless one attends to the conditions under which the rejection of the null hypothesis is *possible*, one can produce nonsignificant findings ad infinitum and any number of such findings will not provide convincing evidence for the truth of the null hypothesis of reasoning–remembering independence.

“LOGICISM” VERSUS OPERATIONAL CONSTRUCTIVISM

Brainerd and Reyna (1992) contrast their fuzzy-trace theory of reasoning and its corollary of reasoning–remembering independence with what they call the “classic” or “logician” position and they identify our arguments with the latter. The “logician” view is described in terms of two propositions: (a) “that memory for critical background information is a necessary precondition for logical reasoning” and, consequently, (b) that “memory independence findings must be due to insensitive experimental conditions that foster reasoning that is not truly logical” (Brainerd & Reyna, 1992, Epilogue). In fact, our position differs substantially from this “logician” view.

Regarding proposition (a), we do agree that memory for *some* background information is necessary for reasoning, but we believe that the information in question depends on the form of reasoning actually used by children in solving the problem. We do *not* agree that the background information in question must be comparisons of adjacent terms, for as we have stated previously (Chapman & Lindemberger, 1988, 1992), children can use a variety of alternative strategies for solving transitivity tasks under suitable conditions. Regarding proposition (b), we do not argue that these alternative strategies are necessarily *illogical*, only that the logic they embody is *different* that that involved in reasoning from adjacent premise comparisons.

For example, compare the inference schemas F1–F3 and T1–T3 given previously. Both schemas are perfectly “logical” in the sense that the third statement in each follows “logically” from the first two statements. The problem is not to determine whether either form of reasoning is “logical” or not, but to identify the logical form in each case. We do not dispute that F1–F3 is “logical,” although we might want to argue that it does not deserve to be called “transitive reasoning” in the usual sense because it does not involve the composition of relations between adjacent

elements which is usually considered the defining characteristic of “transitive reasoning” as such.

Brainerd and Reyna argue for the “logical” character of inference schema F1–F3 by claiming that *information is integrated across the premises* (i.e., across adjacent comparisons) in this inference schema just as it is in T1–T3. We believe this concept of “information integration” is considerably more global than what we have called the “composition of operations” and that focusing on the former tends to blur the otherwise clear distinction between schemas F1–F3 and T1–T3. Consider how children might arrive at the premise F1 (“Things get bigger toward the right”). Presumably, they do it through an *inductive generalization* from particular pairwise comparisons:

C1. Stick A is bigger than stick B, and it’s also on the right of B.

C2. Stick B is bigger than stick C, and it’s also on the right of B.

F1. Things get bigger toward the right.

We believe that such inductive generalizations are ubiquitous in young children’s thinking and that, *under favorable conditions*, they indeed can solve “transitivity tasks” through a generalization such as C1–F1, followed by a deduction of the form F1–F3. We agree that the induction C1–F1 involves an “integration of information” across the individual comparisons C1–C2, but this “integration” is of a fundamentally different type than that occurring in a synthetic inference schema such as T1–T3, in which the conclusion T3 is deduced directly from the individual comparisons T1 and T2. By making “information integration” the sole conceptual criterion for “logical” inference and by failing to differentiate among qualitatively different types of integration, Brainerd and Reyna confound the fundamental distinction between induction and deduction and the corresponding competencies.

In appropriate circumstances, children may succeed in generating a correct answer, either by following the sequence C1–F1, F1–F3 or by means of the schema T1–T3, but the competencies employed are qualitatively different in each case. In particular, T1–T3 involves an understanding of the inferential *relation* between the adjacent comparisons T1–T2 and the conclusion T3, and that is why one would expect the latter to be “dependent” (deterministically!) on the former. In contrast, no such understanding need be involved in following the sequence C1–F1, F1–F3 because the movement from the adjacent comparisons C1–C2 to the conclusion F3 is mediated by the generalization F1. Therefore, there is no reason to expect any dependency (stochastic or otherwise) between F3 and memory for C1–C2.

To repeat the overall argument: One would not expect any dependence between judgments and memory for adjacent comparisons as long as children are reasoning according to the schemas C1–F1, F1–F3, because

the individual comparisons C1–C2 are not the premises of the inference resulting in the judgment F3. Therefore, one cannot infer a general “reasoning–remembering independence” from the lack of an association between memory for C1 and C2 and the judgment F3. On the contrary, a statistically significant stochastic dependence between judgments and memory for individual comparisons would be expected only if (a) the task situation does not allow reasoning according to the schema C1–F1, F1–F3 and (b) sufficient numbers of children use the schema T1–T3 relative to their levels of judgment and memory. Once again, such a result was precisely what was obtained (Chapman & Lindenberg, 1992, Tables 5 and 7).

SPECIFIC CRITICISMS

Once the foregoing confusions are clarified, one can address more easily Brainerd and Reyna’s (1992) specific criticisms: (a) that existing data show reasoning–remembering independence even when the correlation of length with spatial position is controlled; (b) that our own data show reasoning–remembering independence when one controls for multiple significance tests; and (c) that we attribute to them a dual-trace model that they have in the meantime rejected.

Controlling for spatial position. Brainerd and Reyna claim that our model of functional reasoning is contradicted by the results of transitivity studies using visual illusions, mixed symmetrical and asymmetrical comparisons, and free measurement procedures. We think that claim begs the question of whether any of the procedures mentioned in fact provide adequate controls for functional reasoning. Without further task analysis, it is not clear what strategies children would use under those conditions, and it is by no means obvious that the procedures in question provide sufficient controls for functional reasoning. For example, why should a mixed series such as

C'1: A is bigger than B and it's to the right of A,

C'2: B is the same size as C,

C'3: C is bigger than D and it's to the right of D

prevent children from making the inductive generalization that

F'1: Things usually get bigger toward the right

and using that generalization to infer that A is (probably) bigger than D since it is to the right of D. The correlation between size and position does not have to be perfect in order to be noticed and employed.

Moreover, there is nothing special about right–left spatial cues. As we indicated earlier, *any* perceptible dimension that is (or tends to be) correlated with length comparisons could be used as a facilitating cue in an inference scheme like F1–F3. For our part, we have described the kind of control that we think is necessary to eliminate functional reasoning,

namely, presenting adjacent pairwise comparisons individually with spatial position counterbalanced. And under those conditions—given a sufficiently powerful test—stochastic dependence between reasoning and remembering has been found, as we have seen. Finally, as our power analysis indicated, controlling for alternative solution strategies is only part of what is required for detecting reasoning—remembering dependencies. One must also have a sufficient proportion of operational reasoners in the sample, relative to their judgment and memory levels.

Multiple significance tests. Brainerd and Reyna dismiss our finding of reasoning—remembering dependence for our three-term length task by claiming that we inflated the Type I error rate above .05 in conducting five tests of association. As a corrective, they recommend “a McNemar test of the pooled conditions.” The claim is false, and the recommended corrective is inappropriate.

As mentioned previously, the association between judgments and memory for premises that we reported for our three-term length task meets the Bonferroni criterion of .01 ($= .05/5$) needed to protect an overall α level of .05. Such a Bonferroni adjustment is a more appropriate way of controlling the overall α level than performing a McNemar test on the pooled conditions. For one thing, the McNemar test is a test of differences between proportions in correlated samples and not a test of association or stochastic dependence (Fleiss, 1981; Siegel & Castellan, 1988). For another, the procedure of pooling conditions in which the proportion of operational reasoners is sufficient for rejecting the null hypothesis of independence with conditions in which that proportion is not sufficient is a guaranteed method for producing nonsignificant results, if carried far enough. The fact that one can produce nonsignificant results at will by pooling conditions in disregard of the statistical requirements for rejecting the null hypothesis puts the multiple nonsignificant outcomes cited by Brainerd and Reyna into perspective.

Fuzzy-trace theory. We were surprised that Brainerd and Reyna thought our account of fuzzy-trace theory was outmoded because we based that account on the most recent statements of the theory available to us (Brainerd and Reyna, 1990; Reyna & Brainerd, 1990). But they know their theory better than we do, and we refer any readers wishing to check our account to the original sources. In any case, we think that the issues remain the same no matter what version of fuzzy-trace theory one considers.

CONCLUSION

The major issues can be summarized as follows:

1. Brainerd and Reyna believe that the quantity of nonsignificant findings reflects an underlying independence between reasoning and remem-

bering. We believe instead that those nonsignificant findings occurred (a) because memory was tested for premises other than those actually used in reasoning and/or (b) because the proportion of operational reasoners in the sample was insufficient to reject the null hypothesis, given the obtained levels of judgment and memory.

2. We found a dependence between reasoning and remembering when conditions (a) and (b) were controlled. Brainerd and Reyna doubt this finding.

3. Brainerd and Reyna argue that schema F1–F3 is quite as “logical” as schema T1–T3 because both involve the integration of information across premises. We do not dispute that F1–F3 is “logical” in some sense, but we argue that the forms of reasoning and the kinds of information integration are qualitatively different in each case.

Brainerd and Reyna ask one to consider what the data *show*. But that is not in doubt; the data generally show a pattern of nonsignificance. The question is what those data *mean*. Do they reflect an underlying independence between reasoning and remembering or a failure to detect the dependence that exists? We leave it to our readers to decide.

REFERENCES

- Brainerd, C. J. (1973). Judgments and explanations as criteria for the presence of cognitive structures. *Psychological Bulletin*, *79*, 172–179.
- Brainerd, C. J., & Kingma, J. (1984). Do children have to remember to reason? A fuzzy-trace theory of transitivity development. *Developmental Review*, *4*, 311–377.
- Brainerd, C. J., & Kingma, J. (1985). On the independence of short-term memory and working memory in cognitive development. *Cognitive Psychology*, *17*, 210–247.
- Brainerd, C. J., & Reyna, V. F. (1990). Gist is the grist: Fuzzy-trace theory and the new intuitionism. *Developmental Review*, *10*, 3–47.
- Brainerd, C. J., & Reyna, V. F. (1992). The memory independence effect: What do the data show? What do the theories claim? *Developmental Review*, *12*, 164–186.
- Chapman, M., & Lindenberg, U. (1988). Functions, operations, and decalage in the development of transitivity. *Developmental Psychology*, *24*, 542–551.
- Chapman, M., & Lindenberg, U. (1992). Transitivity judgments, memory of premises, and models of children's reasoning. *Developmental Review*, *12*, 124–163.
- Fliess, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed.). New York: Wiley.
- Reyna, V. F., & Brainerd, C. J. (1990). Fuzzy processing in transitivity development. *Annals of Operations Research*, *23*, 37–63.
- Siegel, S., & Castellan, N. J., Jr. (1988). *Nonparametric statistics for the behavioral sciences* (2nd ed.). New York: McGraw-Hill.
- Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society, Series B*, *13*, 238–241.

RECEIVED: December 24, 1990.