May 1998

Collaborative Reasoning: Evidence for Collective Rationality

David Moshman
University of Nebraska-Lincoln, dmoshman1@unl.edu

Molly Geil
University of Nebraska-Lincoln

Correspondence: David Moshman, Department of Educational Psychology, University of Nebraska–Lincoln, Lincoln, NE 68588-0345, USA. Email: dmoshman@unlinfo.unl.edu

Collaborative Reasoning: Evidence for Collective Rationality

David Moshman and Molly Geil

University of Nebraska–Lincoln

Reasoning may be defined as a deliberate effort to coordinate inferences so as to reach justifiable conclusions. Thus defined, reasoning includes collaborative as well as individual forms of cognitive action. The purpose of the present study was to demonstrate a circumstance in which collaborative reasoning is qualitatively superior to individual reasoning. The selection task, a well known logical hypothesis-testing problem, was presented to 143 college undergraduates—32 individuals and 20 groups of 5 or 6 interacting peers. The correct (falsification) response pattern was selected by only 9% of the individuals but by 75% of the groups. The superior performance of the groups was due to collaborative reasoning rather than to imitation or peer pressure. Groups typically co-constructed a structure of arguments qualitatively more sophisticated than that generated by most individuals. The results support Piagetian and Habermasian views of peer interaction as a locus of rational social processes.

INTRODUCTION

Reasoning may be defined as a deliberate effort to coordinate inferences in such a way as to reach justifiable conclusions (Moshman, 1995, 1998). Typically, we think of reasoning as a cognitive action performed by an individual. As defined here, however, reasoning may also be a social process. We may use the term *collaborative reasoning* for cases in which two or more individuals deliberately coordinate their thinking for the shared purpose of achieving justifiable results.

A key aspect of reasoning—individual or collaborative—is testing hypotheses against data (Koslowski, 1996; Kuhn, Amsel, & O'Loughlin, 1988; Moshman, 1998). Good hypothesis-testing requires, among other things, insight into the role of falsification: In order to genuinely test an hypothesis one must seek evidence that would falsify that hypothesis rather than simply accumulate evi-

dence consistent with it. A great deal of recent research has therefore focused on the extent to which people have this basic insight concerning the logic of hypothesis-testing and apply it in testing hypotheses.

Much of the research directly addressing the use of a falsification strategy in hypothesis-testing has involved the selection task, in which individuals are presented with a hypothesis concerning a set of four cards and are asked to decide which cards must be examined to determine whether the hypothesis is true or false. In the original version of the task (Wason, 1966), the hypothesis proposes an arbitrary connection between letters and numbers appearing on the cards. It is well established that, in this version, very few college students select for examination those cards—and only those cards—that could falsify the hypothesis (Green, 1995; Griggs, 1995; Wason & Johnson-Laird, 1972).

Research has shown that the frequency of correct selections can be substantially enhanced through variations on the standard instructions, content, and/or context of the selection task. Not all variations yield better performance, however, and there has been considerable dispute about the explanation for selection task behavior under various circumstances (Evans, 1989; Newstead & Evans, 1995). Some researchers and theorists have concluded that people lack any sort of abstract falsification insight. To the extent that people test hypotheses successfully in certain circumstances it is through the use of domain-specific pragmatic reasoning schemata (Cheng & Holyoak, 1985; Cheng, Holyoak, Nisbett, & Oliver, 1986), concrete mental models (Johnson-Laird & Byrne, 1991), or nonrational heuristics (Evans, 1989; Sperber, Cara, & Girotto, 1995). Others argue, following Piaget (1972), that an abstract understanding of the role of falsification in hypothesis-testing develops over the course of adolescence but that strategic application of that competence depends on a variety of situational and task variables (Bady, 1979; Moshman, 1979; Overton, 1990; Overton et al., 1987; Ward & Overton, 1990).

Notwithstanding their differences, these various theorists share a focus on the competence (or incompetence) of the individual. Correspondingly, research on the selection task, like most psychological research on reasoning, has focused on the performance of individuals working alone. In everyday life, however, a great deal of reasoning is collaborative. Although groups are subject to a variety of biases and often show levels of cognitive performance equal or inferior to the performance of their average members (Gigone & Hastie, 1997; Hinsz, Tindale, & Vollrath, 1997; Kerr, MacCoun, & Kramer, 1996), research on collaborative learning, problem-solving, and decision-making has provided convincing evidence for the cognitive value of peer interaction (Azmitia & Montgomery, 1993; Cohen, 1994; Dimant & Bearison, 1991; Doise & Mugny, 1984; Kruger, 1992, 1993; Moscovici & Doise, 1994; Phelps & Damon, 1989). One may thus wonder whether groups of people reasoning together would show better performance on the selection task than isolated individuals.

Over the past two decades, the first author has presented dozens of undergraduate and graduate classes with the original version of the selection task, recorded students' initial responses on the blackboard, and then asked them to discuss the task with each other and reach a consensus solution. Although few students initially choose the correct (falsification) pattern, it is usually the solution adopted by the class. The decision process, moreover, is highly rational: individuals present a variety of views about which cards should or should not be turned, they present arguments for and against these various views, and typically, over the course of 20–40 minutes, they increasingly comprehend and adopt the falsificationist logic underlying the logically appropriate selections.

The classes thus appear to achieve a qualitatively higher level of performance than individual students. That is, most classes agree to turn those cards, and only those cards, that could falsify the hypothesis, demonstrating a level of metalogical understanding that is rarely seen in individuals responding to the same task. The purpose of the present research was to replicate these classroom observations under more controlled conditions.

METHOD

Subjects

Participants were 143 students from undergraduate educational psychology courses, about equally divided between women and men. Students participated in the study as one means of fulfilling a class requirement.

Materials

Students were asked to complete the original version of the selection task (Wason, 1966). They were presented with four cards pictured on a sheet of paper. Each card had either a letter or a number on it. The letter E was shown on the first card, K on the second card, 4 on the third card, and 7 on the fourth card. It was indicated that each card pictured had a letter on one side and a number on the other. Below the letters was the following hypothesis: "If a card has a vowel on one side then it has an even number on the other side." Students were asked to test the hypothesis by selecting the card(s) they would need to turn over to determine conclusively whether the hypothesis was true or false for the set of four cards shown. At the bottom of the page students were asked to justify their solution in writing, a requirement that, by itself, does not typically improve performance (Griggs, 1995).

Design

Students were randomly assigned to one of three experimental conditions: (a) an individual control condition, (b) an interactive group, or (c) an individual/ inter-

active group. The 32 students assigned to the individual condition were asked to solve the selection task individually. The 54 students assigned to the interactive condition addressed the task collectively in 10 groups of 5 or 6 members each. The 57 students assigned to the individual/interactive condition considered the task first individually and then in 10 groups of 5 or 6 members each.

Procedure

All sessions were conducted by the second author. Prior to beginning the task, students were read a set of general instructions informing them that the purpose of the study was to investigate problem-solving. Students were told they would be required to justify their response to the problem in writing. Students were also encouraged to spend enough time on the task so that they were content with their solution.

After receiving general instructions, students were given a set of instructions for the "four card task." The experimenter emphasized that "all cards have a letter on one side and a number on the other."

During the problem-solving activity the experimenter left the room and watched through a one-way mirror. All sessions were videotaped. Total time to complete the task was recorded for each individual and each group. If a group did not reach consensus after 25 minutes, the experimenter returned and told the group they had 10 minutes to reach consensus. Only one group was unable to reach consensus after 35 minutes, at which time the experimenter intervened and told the students to justify their individual responses for the remaining 10 minutes. (During that time, it turned out, a consensus was reached.)

Students in the individual condition recorded their selections and provided written justifications. They were told that they should take as much time as necessary to make a justifiable selection, that they must participate in the study for the allotted 50 minutes to earn class credit, and that they were not expected to get beyond the initial task. After making and justifying his or her selections, each student was given the correct solution and was asked to justify it in writing.

Members of each interactive group were directed by the experimenter to work together to solve the problem. After the group reached what appeared to be a unanimous judgment, students recorded and explained the solution individually. They were told that they could continue to discuss their conclusions if group members wished.

Members of each individual/interactive group were directed that they should first solve the problem individually and explain their selections in writing, and then discuss their selections with each other in order to achieve a consensus. After the group reached what appeared to be a unanimous judgment, students recorded and explained the solution individually. They were told that they could continue to discuss their conclusions if group members wished.

RESULTS

To comprehend the logic of the selection task one must realize that only a *p* together with a *not-q* falsifies a hypothesis of the form *If p then q*. Thus the cards to be turned are the *p,* which could have a *not-q* on the other side, and the *not-q,* which could have a *p.* For the hypothesis "If vowel then even" and the present cards, then, one must turn the E, which would falsify the hypothesis if it had an odd number on the other side, and the 7, which would falsify the hypothesis if it had a vowel on the other side.

We focus first on the cards selected by individuals and groups. To discern how peer interaction affected individual responses, data from the individual/ interactive groups are then examined in detail.

Individuals vs. Groups

In the individual condition, 9.4% selected the correct *p and not-q* combination. In the interactive condition, 70% of the groups selected the correct *p and not-q* combination; in the individual/interactive condition, 80% of the groups selected this combination. Thus, correct selection patterns were far more common in the group conditions than in the individual condition, χ^2 (1, N = 52) = 23.42, *P* < .001. The difference between interactive and individual/interactive groups was not significant, χ^2 (1, N = 20) = 0.27, *P* > .05. The card combinations selected by students in all conditions are shown in Table 1.

TABLE 1
Patterns Selected by Individuals and Groups

	Condition		
Selection	*Individual Pattern*	*Interactive Group*	*Individual/Interactive Group*
p$\bar{q}$	3	7	8
pq$\bar{q}$	0	0	1
p	5	2	1
p q	14	0	0
All cards	2	1	0
Other	8	0	0
Total	32	10	10

The "‾" above a letter indicates "not." The correct selection is p$\bar{q}$. No pattern in the "other" category was selected by more than two individuals or by any group.

Individual Responses in the Individual/Interactive Condition

Responses of students in the individual/interactive condition were analyzed with respect to any change between the initial and final pattern selected by group members. A change in reasoning was regarded as positive if an individual's initial response was incorrect (any card combination other than *p and not-q*) but his or her final response was correct. If the initial response was correct but the final response was incorrect, the change was considered negative. Table 2 shows the initial and final responses for all individuals in the individual/interactive condition.

Of 57 students, 37 had a change in the correctness of their pattern selection after collaborating with others in their group (i.e. the falsification pattern was their initial or final selection but not both). Of these 37, 35 students showed positive change; only two students changed in a negative direction (binomial *P* < .001). Thus, most individuals showed a change in the correctness of their selection as a result of group interaction and almost all of these changes were in a positive direction.

Individual/Interactive Group Patterns

Because individuals in the individual/interactive condition responded to the task prior to peer interaction, it was possible to relate group decisions to initial individual responses. Three of the individual/interactive groups initially contained a majority response. In each of these groups the *p and q* combination was the majority (but not unanimous) initial response. Two of these three groups switched

TABLE 2

Initial and Final Patterns Selected by Members of
Individual/Interactive Groups

Initial Selection Pattern	*Final Selection Pattern*			
	$p\bar{q}$	$pq\bar{q}$	p	*Total*
$p\bar{q}$	10	1	1	12
$pq\bar{q}$	2	0	0	2
p	4	0	2	6
$p\ q$	21	4	2	27
All cards	4	1	0	5
$p\ p$	3	0	1	4
p	1	0	0	1
Total	45	6	6	57

The "¯" above a letter indicates "not." The correct selection is $p\bar{q}$.

to the correct pattern after collaboration. In three other groups, although there was no majority response, *p and q* was the modal (i.e. most frequent) initial response. Again, two of these groups selected the correct pattern after collaboration. In one group the modal response was to turn all four cards, but after collaborating the group selected the correct pattern. The remaining three groups initially contained no modal response. All three of these groups selected the correct *p and not-q* combination after collaboration.

Individual/interactive groups also varied with respect to whether or not they contained individuals who selected the correct pattern initially. Seven of the ten groups had at least one member who initially selected the correct *p and not-q* combination. In no case was this the modal response. Nonetheless, five of these groups selected the correct pattern after collaboration; only two did not. In the remaining three groups, none of the members initially selected the correct response. Nevertheless, each of these groups ultimately agreed on the correct *p and not-q* combination. Thus, having a member who initially selected the correct cards was neither a necessary nor sufficient condition for group success.

Examples of Collaborative Reasoning

With rare exceptions involving developmental progress in children (Doise & Mugny, 1984), research suggests that group performance on cognitive tasks is typically within the range of individual performance (Gigone & Hastie, 1997; Kerr et al., 1996). At the upper limit, on tasks with demonstrably correct solutions, a group may perform at the level of its best member (Laughlin & Ellis, 1986). Presumably, the individual who best comprehends the logic of the task is able to convince the others of his or her view. Of particular interest in the present study, then, are the three groups in which no member initially made the correct selection. In order to illustrate the process of collaborative reasoning, we present summaries of the discussion for two of these groups (the videotape of the third was accidentally destroyed).

Group 1 (clockwise around the table) consisted of Alice, Ben, Carol, Dan, and Earl (the names are fictitious but Alice and Carol were female and the others male). In their initial individual selections, everyone but Alice proposed turning only the E and 4 cards. Alice proposed E, 4, and 7.

At the onset of discussion, everyone immediately agrees that E should be turned to see if it has an even number on the other side, 4 should be turned to see if it has a vowel, and the other two cards need not be turned because the hypothesis does not address them. Alice notes that she had proposed turning E, 4, and 7 but dismisses her selection of 7 as simply a mistake. "I was thinking that it was even for some strange reason," she laughs. "I wasn't thinking, that was it!"

As the discussion proceeds, however, with Alice, Ben, Carol, and Dan rein-forcing each other's statements as to the relevance of E and 4 to the hypothesis and the irrelevance of K and 7 to the hypothesis, Alice suddenly seems to take her initial selection of 7 more seriously:

Alice:	Maybe 7 has a vowel on the other side.
Ben:	It could, but as far as this hypothesis here it just doesn't matter.
Alice:	But if it has…
Dan:	It just says if it has a vowel on one side.
Alice:	Yeah, but it says if it has a vowel on one side, then it has an even number on the other side.
Dan:	So maybe we're wrong.
Carol	[surprised and excited]: Oh that's *true*!
Alice:	7 *could* have a vowel on the other side.

Everyone agrees that the issue is more complex than it initially appeared. Ben proposes that turning the E and 4 would "test" the hypothesis and finding an even number and a vowel respectively on the reverse of those cards would "support it as opposed to proving it," but suggests that "we couldn't prove it un-less we turned over all of them." After substantial discussion about the mechan-ics of turning four cards simultaneously and keeping track of which is which, Dan abruptly focuses the conversation on K:

Dan:	Do we need to turn over K?
Alice:	I don't think so because …
Dan:	We don't have to turn over K.
Alice:	We're concerned with vowels.
Dan:	It has a letter and a number and we know that that one [the K] has a letter and it's not a vowel.
Alice:	Yeah.
Carol:	But what if it has an even number on the other side?
Dan:	It doesn't say anything about …
Alice:	It doesn't say that if it's a consonant …
Dan:	It just says if it has a vowel. It doesn't say if it has a consonant it can't have any of them.
Carol:	That's true.
Dan:	I don't think we have to turn over K.

There is a pause in the conversation until Carol, apparently convinced that turning K is not necessary and now wondering whether 4 is also unnecessary, adds "and it doesn't say that if it has an even number on one side it has a vowel on the other." After another pause she adds, "Really we don't need to turn over 4." She and Ben expand on this as follows:

Carol:	You don't have to turn over 4 because it [the hypothesis] says if it has a vowel on one side it has an even number on the other side.

Ben: It doesn't say if it has an even number on one side it has a vowel on
             the other.

Alice and Dan continue to insist that the 4 must be turned, however. After
discussion among all four, Carol starts to explain that the 4 "could have a con-
sonant on the other side and it still wouldn't …" but Alice interjects, "Yeah, but
we need to check it because it *is* an even number, so we have to find out if it *has*
a vowel on the other side." Carol reluctantly concurs: "Yeah, I guess."

Alice then continues, explaining that, in addition to 4, the E must be turned
because "we have to find out if it's an even number" and the 7 must be turned
because "we need to find out if it's a consonant or a vowel because if it's a
vowel then it's false." Ben and Carol express continuing uncertainty about the
4, however. Ben wonders about the implications of it having a consonant on the
reverse side, leading Carol to observe that turning the 4 would "prove" the hy-
pothesis if there were a vowel on the reverse side and "wouldn't do anything"
if there were a consonant. "It would either prove it or it wouldn't do anything,"
she concludes. "The 7 and the E are the only ones that can disprove it."

There is a very long pause, during which Alice, looking back at the instruc-
tions, can be heard murmuring, "to determine conclusively." Finally, Dan gets
the discussion going again and, with input from everyone, a new consensus is
reached:

Dan: Okay, we have to turn over E for sure, right?
Carol: Yeah.
Dan: Because it has a vowel on one side and we need to find out if it has an
             even number on the other. K we don't have to worry about because it
             doesn't say anything about …
Ben: … consonants …
Dan: … having a consonant.
Alice: It doesn't say if it has a consonant it's odd or whatever.
Dan: And 4 …
Ben: I think we need to turn 4.
Dan: I think we have to turn over 4 because …
Earl: It's the same as E, really.
Alice: It's the same as E, yeah, we know it's an even number so we have to
             find out if it has …
Dan: Well maybe we don't. [He pauses, then proceeds slowly, with Carol
             nodding and murmuring assent.] If it has a consonant on one side it
             doesn't matter if it has an odd or an even number. So it really doesn't
             matter if we look at 4. Does it? Do you see that?
Alice and Ben (simultaneously): I see what you're saying.
Dan: It can tell us where that's right, but it can't tell us it's wrong.
Carol: Yeah.
Dan: And 7 I think we have to turn over. 'Cause we need to find out if that
             has a vowel.

| Carol: | Because it can prove right or wrong. |
| Dan: | Because it can prove it wrong. |

The group then confirms that consensus has been reached:

Ben:	So are we narrowing it down to E and 7 this time?
Dan:	I think so.
Carol:	I think it should be E and 7 now.
Dan:	I do too.
Ben:	That's pretty interesting to watch us all concur.
Alice:	I wouldn't have come up with this if we hadn't, you know, talked about it.
Carol:	I know, I was totally set on E and 4.
Ben:	We *all* were.

After brief additional discussion reviewing the irrelevance of 4 and the irreversibility of the hypothesis, each of the five students independently wrote on his or her final task sheet that only the E and 7 should be turned and each provided a written explanation consistent with the group's final arguments.

Group 2 (clockwise around the table) consisted of Frank, Gwen, Hal, Ike, and Jay (again the names are fictitious but Gwen is female and the others male). In their initial individual selections, Hal proposed to turn only the E, Frank proposed E, 4, and 7, and the other three proposed E and 4.

The discussion begins with Jay suggesting and Gwen agreeing that E and 4 are the cards that must be turned. Frank also agrees that E and 4 must be turned but notes that he selected 7 as well. Reviewing his reasoning, he explains that E must be turned because it is a vowel and K need not be turned "because we're not concerned with a consonant having an odd or even; we're talking about just the vowel card." With respect to the other side of the 4, he starts to explain, "if that's a vowel, then …" Hal interrupts: "See, that doesn't have to be a vowel, though." Gwen objects, "Yeah, it does," but Hal goes on to explain his view, concluding, with respect to the hypothesis:

All that that's saying is that all the vowels have to have even numbers on the other side. But it's not saying that like the number 4 has to have a vowel on the other side.

Frank and Gwen agree now, and Hal goes on:

The only thing that you have to turn over is E, but by turning over E you're just supporting the hypothesis, you're not proving. The only way you can prove that thing right is if you have all the vowels facing up and then you turn all of them over. You know what I'm saying? That's the only way you can prove that hypothesis correct.

Frank now asserts "you'd have to turn over 7." Hal responds, "you'd have to turn over all of the vowels and they'd have to be facing up too. You couldn't have one facing the other way because you wouldn't know." Gwen and Jay agree; further discussion among Frank, Gwen, and Hal addresses the requirements of the task, the need to turn E, and the irrelevance of K and 4. Frank then returns to the question of the 7 card:

Frank:	But we would have to turn over the 7, then. Because a vowel like an I or
	an O would throw us off.
Gwen:	You're right, we wouldn't be sure of the 7.
Frank:	So, the E and the 7?
Gwen:	Yeah.
Hal:	Yeah, that's a good point.
Gwen:	It *is* a good point. [Pause.] So we've decided E and 7 because we want to know if E has an even number under it and if 7 has a vowel under it or a consonant.
Frank:	That would disprove it.
Gwen:	Right. That's interesting.

Frank goes on to explain again that "we do not need to turn over card K because our hypothesis is not concerned with a consonant having an odd or even number," and then he and Gwen summarize the case for choosing 7 as follows:

Frank:	If card 7 is a consonant, our hypothesis, no, if card 7 is a *vowel* …
Gwen:	…vowel, then the hypothesis is false.
Frank:	Right.

There appeared to be general assent at this point and each student then indicated the choice of E and 7 on his or her final task sheet and provided an explanation consistent with the group's final arguments.

DISCUSSION

It appears that, under the conditions of the present research, most groups, but few individuals, applied a falsification strategy in testing a hypothesis. Groups, in other words, typically acted in accord with logical norms of hypothesis-testing, whereas individuals did not. Results from the individual/interactive condition, moreover, indicate achievement of a genuine consensus: after participating in group discussion of the selection task, individuals usually rejected their initial response in favor of the falsification pattern. Insight into the logic of falsification, it appears, is more readily achieved in collaborative reasoning than in individual reasoning.

Our discussion of these results addresses three general issues. First, we consider and reject the possibility that the evidence for superior group performance

is an artifact of our methodology or the result of nonrational social influences. Second, we examine the nature of collective rationality, concluding that the groups in our study engaged in a process of argumentation involving the co-construction of a collective "reckoning." Finally, we consider the relation of collective rationality to individual rationality.

Nonrational Influences on Group Behavior

A variety of factors may be invoked in explaining the present evidence. Roughly, two general categories of explanation may be differentiated. One type of account for the superiority of group performance and the change in individual response patterns during peer interaction would emphasize factors such as time on task, individual influence, or concurrence-seeking behavior (Russell, Mills, & Reiff-Musgrove, 1990; Veiga, 1991). An alternative approach to the role of peer interaction highlights rational group processes including clear formulation and communication of ideas, critical evaluation of proposed judgments and choices, reflection on assumptions and perspectives, and free discussion of alternatives (Azmitia & Montgomery, 1993; Binkley, 1995; Dimant & Bearison, 1991; Doise & Mugny, 1984; Goldman, 1994; Habermas, 1990; Kruger, 1992, 1993; Meloth & Deering, 1994; Moscovici & Doise, 1994; Moshman, 1995, 1998; Phelps & Damon, 1989; Piaget, 1965, 1995; Salmon & Zeitz, 1995; Slade, 1995; Wheatley, 1991; Youniss & Damon, 1992).

With respect to nonrational variables, one potential explanation for the superiority of group performance is that groups simply took more time to evaluate their selection patterns before agreeing on a response. Groups did indeed, on average, take about two and a half times as long as individuals to make a selection. Students in the individual condition, however, were encouraged to take as much time as necessary to make a selection they were content with and could justify in writing. Moreover, they were told they must participate in the study for the allotted 50 minutes to receive class credit, and that they were not expected to get beyond the task. Thus, they were given every reason to take their time.

When students in the individual condition finished responding, they were given the correct solution to the task and asked to justify it. If time were a significant factor in outcome, then giving students the right answer and encouraging them to further reflect on the task might have improved understanding. There was little evidence of such improvement, however. In fact, 49% of the students who were asked to justify the correct response pattern either argued that the *p and not-q* combination was incorrect or indicated that they did not know why you would turn those cards. Most of the remaining 51% gave ambiguous justifications. Evans and Wason (1976) found that their students were more likely than ours to accept and justify the correct solution when it was presented

to them, but found that incorrect solutions were also accepted and rationalized. It appears unlikely, in any event, that the effects of peer interaction were due simply to the fact that groups spent more time on the task than individuals.

Another possible explanation for group performance and changes in individual response patterns is conformity to peer pressure. The response patterns in the individual/interactive condition do not support this explanation, however. In none of those 10 groups was the correct pattern chosen by a majority of individuals prior to the discussion; in fact, it was never even the most common pattern. Thus the dramatic increase in correct responding associated with group discussion was not a matter of conformity to the initial majority or modal view.

It might be suggested that even if the correct pattern was not initially most common, those arguing for it were individuals with greater social influence. The participants in the study, however, were all randomly assigned undergraduates; there was no reason why those arguing for any particular response would be seen, independent of the quality of their arguments, as more knowledgeable or worthy of emulation than those arguing for any other response. It is noteworthy, moreover, that in three of the ten individual/interactive groups not a single student initially chose the correct response. Agreement on the correct response was nonetheless reached in all three of these groups. It seems likely, then, that social interaction increased correct responding to the selection task not by coercing conformity but by facilitating genuine insight into logical aspects of hypothesis-testing.

Collective Rationality as Co-constructive Argumentation

In every individual/interactive group, the initial selections included at least two different responses to the task. Videotapes of the discussions show that students routinely challenged each other to justify their selections and encouraged each other to consider consequences and alternatives. As illustrated earlier, final selections generally appeared to reflect voluntary agreement based on genuine insight into the logic of the task. Our tapes reveal little evidence of passive conformity to majority views or to the views of an apparent expert. On the contrary, we routinely see students attempting to reach a genuine consensus by sharing insights, perspectives, reasons, doubts, and possibilities.

Research on group cognitive performance shows that on tasks with demonstrably correct solutions groups frequently perform at the level of their best member (Gigone & Hastie, 1997; Kerr et al., 1996; Laughlin & Ellis, 1986). This suggests that collective rationality may be a matter of rational group choice. Given two or more potential solutions to a problem and arguments for and against these various solutions, groups are likely to choose the most defensible solution rather than the solution initially favored by the largest number of group members.

Such a conception, however, is insufficient to explain our results. In particular, it cannot account for the three groups in our study that reached consensus on the correct selection pattern despite the fact that no member of the group initially selected this pattern. We thus consider again the nature of the discussions in the two groups presented in detail earlier.

In Group 1, four of the five students are highly involved in formulating the group response. Alice is the only student who initially proposed turning the 7, and is the first to note that it could have a vowel on the reverse side. Ben is the first to make a distinction between "supporting" the hypothesis and "proving" it, but mistakenly concludes that all four cards must be turned to prove the hypothesis. Dan then takes the lead in arguing that, notwithstanding Ben's point, there is no need to turn the K. Carol then expands on Dan's arguments to make the case that there is also no need to turn the 4, initially convincing Ben but not Alice or Dan. Alice then notes that the 7 could falsify the hypothesis. Later, when the conversation returns to 4, Carol notes that the 4 could not disprove the hypothesis whereas the E and 7 could. It is Dan, however, who later establishes that the group has agreed to turn 7 and who provides what turns out to be the decisive argument for not turning 4.

In Group 2, three of the five students play major roles in the discussion. Frank is a persistent advocate for turning the 7, finally convincing the others. Hal makes the case that the 4 need not be turned, though initially in the context of arguing that one need only turn E. Gwen initially insists that 4 must be turned but is convinced by Hal's argument against this. She later is the first to appreciate the significance of Frank's argument for turning 7 and plays a key role in helping him clarify it and convince the others.

It appears, then, that at least seven of the ten students in these two groups were actively engaged in a process of collaborative reasoning that generated a higher level of understanding than any of them alone could muster. Although the process included rational choices among conflicting alternatives, the alternatives were not complete solutions but component ideas. Group 2, for example, ultimately accepted Frank's argument for turning the 7 and Hal's argument for not turning the 4, although each was initially alone in his choice. It is notable, however, that the group integrated these two insights to generate the correct solution despite the fact that neither Frank, Hal, nor anyone else had initially proposed that solution.

The process in Group 1 is even more interesting. Here too, the group concluded that 7 must be turned, despite the fact that only one member initially made this selection. It is particularly noteworthy, however, that the group also agreed that the 4 need not be turned, despite the fact that all five members initially thought it should. Moreover, it is impossible to identify a particular individual whose insights and arguments are the basis for the group consensus. Carol is the first to propose that 4 should not be turned, but her arguments draw on (a) Dan's case against turning K; (b) an earlier distinction by Ben between

"supporting" and "proving" the hypothesis; and (c) Alice's insight, with regard to 7, about the central relevance of falsification. It is Dan, moreover, who ultimately, with assistance from Carol, convinces the others that 4 need not be turned. The group's decision not to turn 4, then, is not a choice among conflicting individual ideas but rather a novel insight co-constructed by four members of the group.

Kruger (1993) raises the question of whether peer interaction with respect to cognitive tasks is best construed as a process of conflict or a process of cooperation. We agree with her conclusion that this is a false dichotomy. There were indeed conflicting views, and choices among these views, within every group in our study. These conflicts, however, took place within a cooperative context in which group members focused not on proving their own views to be correct, but on co-constructing a consensus solution that could be justified to everyone's satisfaction. The result for each group was not just a pattern of choices but a structure of arguments justifying those choices, a collective version of what Binkley (1995) labels a "reckoning." The higher level of understanding achieved by most groups, then, may be conceptualized as a well elaborated collective reckoning, in this case a structure of arguments supporting the selection of E and 7 and explaining the irrelevance of K and 4.

Individual and Collective Rationality

Although the primary focus of the present research was on collaborative reasoning, it is not our intent to downplay the importance of individual reasoning or to sharply distinguish individual and collective rationality. Green (1995) argues convincingly that individual thinking about the selection task involves a process of internal argumentation. Developmental psychologists, moreover, increasingly construe individual cognition not as a single equilibrated structure but rather as an interplay of competing ideas, strategies, and perspectives (Kuhn, Garcia-Mila, Zohar, & Andersen, 1995; Siegler, 1996). The coordination of competing arguments to construct a novel reckoning may thus be typical of individual as well as collaborative reasoning (cf. Kuhn, 1991). In the case of the selection task, however, it appears from our data that groups are more likely than individuals to construct the complete reckoning that supports the correct selection of cards. Most groups thus achieve a level of insight qualitatively superior to that achieved by most individuals. The process of collaborative reasoning that generates this higher level of insight, however, may not be qualitatively different from individual processes of reasoning.

Group performance is no doubt constrained by the cognitive competencies of the individuals constituting the group. In the present case, for example, the success of the groups may depend on the same metalogical understanding of falsification that enables individual success by adolescents and adults on other versions of the selection task (Bady, 1979; Moshman, 1979; Overton, 1990;

Overton et al., 1987; Ward & Overton, 1990). If so, one might expect that peer groups of children under age 11 or 12 would fare no better than individual children on the present task.

Rather than view individual competence as a given, however, it is important to note that participation in collaborative reasoning may facilitate individual cognitive progress. Piaget, Habermas, and other theorists have identified peer interaction as a locus of rational social processes (Binkley, 1995; Goldman, 1994; Habermas, 1990; Moscovici & Doise, 1994; Piaget, 1995; Salmon & Zeitz, 1995) that promote developmental change (Moshman, 1995, 1998; Piaget, 1965; Youniss & Damon, 1992). Psychological research has confirmed that individuals may progress towards higher levels of cognition as a result of peer interaction (Azmitia & Montgomery, 1993; Dimant & Bearison, 1991; Doise & Mugny, 1984; Kruger, 1992, 1993). Peer interaction is, moreover, widely regarded as a highly effective basis for education (Cohen, 1994; Meloth & Deering, 1994; Phelps & Damon, 1989; Slade, 1995; Wheatley, 1991).

The present results suggest that levels of understanding that are difficult to elicit in situations of individual performance may emerge in the context of collaborative reasoning. We are not proposing that the brief peer interaction experienced by the participants in the present study produced a generalizable and durable change in individual hypothesis-testing competence. It is nevertheless plausible that repeated opportunities to participate in sophisticated group reasoning may contribute to the development of individual reasoning, which in turn enhances the quality of one's epistemic transactions with others. Human rationality may thus develop via an ongoing dialectic of individual and collaborative reasoning.

REFERENCES

Azmitia, M., & Montgomery, R. (1993). Friendship, transactive dialogues, and the development of scientific reasoning. *Social Development*, *2*, 202–221.

Bady, R.J. (1979). Students' understanding of the logic of hypothesis testing. *Journal of Research in Science Teaching*, *16*, 61–65.

Binkley, R.W. (1995). Argumentation, education, and reasoning. *Informal Logic*, *17*, 127–143.

Cheng, P.W., & Holyoak, K.J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, *17*, 391–416.

Cheng, P.W., & Holyoak, K.J., Nisbett, R.E., & Oliver, L.M. (1986). Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive Psychology*, *18*, 293–328.

Cohen, E.G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research*, *64*, 1–35.

Dimant, R.J., & Bearison, D.J. (1991). Development of formal reasoning during successive peer interactions. *Developmental Psychology*, *27*, 277–284.

Doise, W., & Mugny, G. (1984). *The social development of the intellect*. Oxford: Pergamon Press.

Evans, J.St.B.T. (1989). *Bias in human reasoning: Causes and consequences*. Hove, UK: Lawrence Erlbaum Associates Ltd.

Evans, J.St.B.T., & Wason, P.C. (1976). Rationalisation in a reasoning task. *British Journal of Psychology*, *63*, 205–212.

Gigone, D., & Hastie, R. (1997). Proper analysis of the accuracy of group judgments. *Psychological Bulletin*, *121*, 149–167.

Goldman, A.I. (1994). Argumentation and social epistemology. *Journal of Philosophy*, *91*, 27–49.

Green, D.W. (1995). The abstract selection task: Thesis, antithesis, and synthesis. In S.E. Newstead & J.St.B.T. Evans (Eds.), *Perspectives on thinking and reasoning: Essays in honor of Peter Wason* (pp. 173–188). Hove, UK: Lawrence Erlbaum Associates Ltd.

Griggs, R.A. (1995). The effects of rule clarification, decision justification, and selection instruction on Wason's abstract selection task. In S.E. Newstead & J.St.B.T. Evans (Eds.), *Perspectives on thinking and reasoning: Essays in honor of Peter Wason* (pp. 17–39). Hove, UK: Lawrence Erlbaum Associates Ltd.

Habermas, J. (1990). *Moral consciousness and communicative action*. Cambridge, MA: MIT Press.

Hinsz, V.B., Tindale, R.S., & Vollrath, D.A. (1997). The emerging conceptualization of groups as information processors. *Psychological Bulletin*, *121*, 43–64.

Johnson-Laird, P.N., & Byrne, R.M.J. (1991). *Deduction*. Hove, UK: Lawrence Erlbaum Associates Ltd.

Kerr, N.L., MacCoun, R.J., & Kramer, G.P. (1996). Bias in judgment: Comparing individuals and groups. *Psychological Review*, *103*, 687–719.

Koslowski, B. (1996). *Theory and evidence: The development of scientific reasoning*. Cambridge, MA: MIT Press.

Kruger, A.C. (1992). The effect of peer and adult–child transactive discussions on moral reasoning. *Merrill-Palmer Quarterly*, *38*, 191–211.

Kruger, A.C. (1993). Peer collaboration: Conflict, cooperation, or both? *Social Development*, *2*, 165–182.

Kuhn, D. (1991). *The skills of argument*. Cambridge, UK: Cambridge University Press.

Kuhn, D., Amsel, E., & O'Loughlin, M. (1988). *The development of scientific thinking skills*. New York: Academic Press.

Kuhn, D., Garcia-Mila, M., Zohar, A., & Andersen, C. (1995). Strategies of knowledge acquisition. *Monographs of the Society for Research in Child Development*, *60*(4), Serial No. 245.

Laughlin, P.R., & Ellis, A.L. (1986). Demonstrability and social combination processes on mathematical intellective tasks. *Journal of Experimental Social Psychology*, *22*, 177–189.

Meloth, M.S., & Deering, P.D. (1994). Task talk and task awareness under different cooperative learning conditions. *American Educational Research Journal*, *31*, 138–165.

Moscovici, S., & Doise, W. (1994). *Conflict and consensus: A general theory of collective decisions*. London: Sage.

Moshman, D. (1979). Development of formal hypothesis-testing ability. *Developmental Psychology*, *15*, 104–112.

Moshman, D. (1995). Reasoning as self-constrained thinking. *Human Development*, *38*, 53–64.

Moshman, D. (1998). Cognitive development beyond childhood. In D. Kuhn & R. Siegler (Eds.), W. Damon (Series Ed.), *Handbook of child psychology (5th edn.): Vol. 2: Cognition, perception, and language.* New York: Wiley

Newstead, S.E., & Evans. J.St.B.T. (Eds.) (1995). *Perspectives on thinking and reasoning: Essays in honor of Peter Wason.* Hove, UK: Lawrence Erlbaum Associates Ltd.

Overton, W.F. (1990). Competence and procedures: Constraints on the development of logical reasoning. In W.F. Overton (Ed.), *Reasoning, necessity, and logic: Developmental perspectives* (pp. 1–32). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Overton, W.F., Ward, S.L., Noveck, I.A., Black, J., & O'Brien, D.P. (1987). Form and content in the development of deductive reasoning. *Developmental Psychology*, *23*, 22–30.

Phelps, E., & Damon, W. (1989). Problem solving with equals: Peer collaboration as a context for learning mathematics and spatial concepts. *Journal of Educational Psychology*, *81*, 639–646.

Piaget, J. (1965). *The moral judgment of the child*. New York: Free Press. [Originally published 1932.]

Piaget, J. (1972). Intellectual evolution from adolescence to adulthood. *Human Development*, *15*, 1–12.

Piaget, J. (1995). *Sociological studies*. London: Routledge. Russell, J., Mills, I., & Reiff-Musgrove, P. (1990). The role of symmetrical and asymmetrical social conflict in cognitive change. *Journal of Experimental Child Psychology*, *49*, 58–78.

Salmon, M.H. & Zeitz, C.M. (1995). Analyzing conversational reasoning. *Informal Logic*, *17*, 1–23. Siegler, R.S. (1996). *Emerging minds: The process of change in children's thinking*. Oxford: Oxford University Press.

Slade, C. (1995). Reflective reasoning in groups. *Informal Logic*, *17*, 223–234.

Sperber, D., Cara, F., & Girotto, V. (1995). Relevance theory explains the selection task. *Cognition*, *57*, 31–95.

Veiga, J.F. (1991). The frequency of self-limiting behavior in groups: A measure and an explanation. *Human Relations*, *44*, 877–895.

Ward, S.L., & Overton, W.F. (1990). Semantic familiarity, relevance, and the development of deductive reasoning. *Developmental Psychology*, *26*, 488–493.

Wason, P.C. (1966). Reasoning. In B.M. Foss (Ed.), *New horizons in psychology*. Harmondsworth, UK: Penguin.

Wason, P.C., & Johnson-Laird, P.N. (1972). *Psychology of reasoning: Structure and content*. Cambridge, MA: Harvard University Press.

Wheatley, G.H. (1991). Constructivist perspectives on science and mathematics learning. *Science Education*, *75*, 9–21.

Youniss, J., & Damon, W. (1992). Social construction in Piaget's theory. In H. Beilin & P.B. Pufall (Eds.), *Piaget's theory: Prospects and possibilities* (pp. 267–286). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.