COGNITIVE COOPERATION

When the Going Gets Tough, Think as a Group

David Sloan Wilson, John J. Timmel, and Ralph R. Miller
Binghamton University

Cooperation can evolve in the context of cognitive activities such as perception, attention, memory, and decision making, in addition to physical activities such as hunting, gathering, warfare, and childcare. The social insects are well known to cooperate on both physical and cognitive tasks, but the idea of cognitive cooperation in humans has not received widespread attention or systematic study. The traditional psychological literature often gives the impression that groups are dysfunctional cognitive units, while evolutionary psychologists have so far studied cognition primarily at the individual level. We present two experiments that demonstrate the superiority of thinking in groups, but only for tasks that are sufficiently challenging to exceed the capacity of individuals. One of the experiments is in a brainstorming format, where advantages of real groups over nominal groups have been notoriously difficult to demonstrate. Cognitive cooperation might often operate beneath conscious awareness and take place without the need for overt training, as evolutionary psychologists have stressed for individual-level cognitive adaptations. In general, cognitive cooperation should be a central subject in human evolutionary psychology, as it already is in the study of the social insects.

KEY WORDS: **Brainstorming; Cooperation; Evolutionary psychology; Group cognition; Group decision-making**

Cooperation is found throughout the animal kingdom and is especially common in our own species. For cooperation to evolve, there must first

Received May 7, 2003; accepted September 18, 2003; final version received December 30, 2003.

Address all correspondence to David Sloan Wilson, Department of Biology, Binghamton University, Binghamton, NY 13902-6000. Email: Dwilson@binghamton.edu

Human Nature, Vol. 15, No. 3, pp. 1-15. **1045-6767/98/$6.00 + .15**

be a task that requires the coordinated action of more than one individual. Then it must be possible to solve the problems of cheating that often accompany coordinated action. Sometimes there is little incentive to cheat because cooperation produces large benefits for everyone at trivial individual cost. At other times cooperation is more costly and evolves only in groups where genetic relatedness is high or social control mechanisms are in place. Social insect colonies are one pinnacle of cooperation in the animal kingdom. Human social groups are another pinnacle, although the evolutionary pathways were not necessarily the same in the two cases (Sober and Wilson 1998).

Cooperation is usually studied in the context of physical activities such as hunting, gathering, warfare, or childcare. However, cooperation can also take place in the context of cognitive activities such as perception, attention, memory, and decision-making. All of these cognitive processes can potentially benefit from the coordinated action of more than one individual, just as physical activities do. Thus, we should expect to see animals merging their minds in addition to their muscles—perhaps even more so if the benefits of cooperation are greater and the problems of cheating less for cognitive than for physical activities.

The idea of cognitive cooperation in addition to physical cooperation is well understood by social insect biologists. For example, a honeybee colony monitors and adaptively responds to its resource environment over an area of several square kilometers. When the quality of a nectar source is artificially raised and lowered, the hive responds within hours with an appropriate allocation of workers (Seeley 1995). Similarly, when a colony splits and a new nest site must be found, scouts that have investigated single sites integrate their information and the swarm moves directly to the best site (Seeley and Buhrman 1999). Just as individual cognition requires neuronal interactions in which each neuron plays a small role in the process, group cognition in social insect colonies requires social interactions in which each insect plays a small role. The idea of a "group mind" might sound like science fiction, but it has been firmly established for the social insects by the meticulous experiments of Seeley and many others (e.g., Camazine et al. 2001).

If social insects can merge their minds in addition to their muscles, how about humans? Strangely, cognitive cooperation is not well studied in our own species. The first social scientists imagined human societies as organic units, complete with "group minds," but this holistic perspective was largely rejected during the middle of the twentieth century (Wegner 1986). The modern psychological literature includes hundreds of papers that employ terms such as "group problem solving" and "group decision making," consisting of numerous specific research programs on various cognitive tasks. It is difficult to summarize such a diverse literature, but the verdict often appears to be that human groups do *not* function well as cognitive units. Two outstanding examples are the concepts of "groupthink" and "brainstorming." Janis (1972, 1982) claimed that groups are prone to faulty decision-making and that these deficits were responsible for foreign policy disasters such as the Bay of Pigs and the Vietnam war. His term "groupthink" became a household word and invokes an image of groups as greatly inferior to individuals as cognitive units. The term "brainstorming" was coined by an advertising executive who claimed that people generate more and better ideas in groups than alone (Osborne 1957). Dozens of studies attempted to confirm this claim by comparing the performance of groups composed of interacting individuals with the performance of an equal number of individuals thinking by themselves (nominal groups) on a variety of tasks. The results were so uniformly negative that Mullen, Johnson, and Salas (1991) concluded their meta-analysis with the following strong statement: "It appears to be particularly difficult to justify brainstorming techniques in terms of any performance outcomes, and the long-lived popularity of brainstorming techniques is unequivocally and substantively misguided." Current research on brainstorming has largely stopped looking for performance advantages of groups and instead attempts to explain the performance deficits (Brown and Paulus 2002; Stroebe and Diehl 1994).

Wilson (1997) has reviewed the traditional psychological literature on group decision making from an evolutionary perspective. His assessment will be summarized in the discussion section of this paper, but first it is necessary to ask what the newer field of evolutionary psychology has to say on the subject. Evolutionary psychology has had much to say about cognition at the individual level, such as the existence of innate specialized adaptations to solve problems encountered in the ancestral environment (Cosmides and Tooby 1992), the need to receive information as frequencies rather than percentages (Gigerenzer and Hoffrage 1995), and the use of simple heuristics that make us efficiently smart rather than trying to comprehensively solve problems (Gigerenzer et al. 1999). However, the possibilities that groups might engage in coordinated cognitive processes, and that some of the insights just listed for individual cognition might also apply to group cognition, have not been considered.

To summarize, the idea of cognitive cooperation in humans is in a highly unsettled state. We are known to be a cooperative species, and cognitive tasks are as amenable to cooperation as physical tasks. The other pinnacle of sociality (the social insects) is known to engage in cognitive cooperation, yet traditional psychologists appear to find little evidence for it in humans and evolutionary psychologists haven't even started looking.

In this paper we show that cognitive cooperation can be demonstrated in humans, even in brainstorming experiments, within which it has previously been notoriously difficult to find. Furthermore, cognitive cooperation might operate beneath conscious awareness and without the need for learning, much as evolutionary psychologists have emphasized for individual-level cognition. Our experiments only begin to address a very large subject, but they suggest that cognitive cooperation should occupy center stage in human evolutionary psychology, as it already does in the study of social insects.

THE EXPERIMENTS

Cooperation is most useful for tasks that exceed the ability of individuals acting alone. Most people can lift a glass of water by themselves, and efforts to help would only get in the way. Most people cannot lift a piano by themselves, and coordinated action is absolutely essential. This point is obvious in the context of physical activities, but it has not been sufficiently appreciated in the context of cognitive activities. For example, we cannot find a single brainstorming experiment that manipulates task difficulty as an independent variable.

We therefore made task difficulty the focus of our research. We chose the game of Twenty Questions because it is familiar, can be played by either individuals or groups, and has a number of advantages for studying task difficulty (Taylor and Faust 1952). The object of the game is to guess a word by asking no more than 20 questions that can be answered with "yes," "no," or "ambiguous." Determining the word out of a set of many possibilities can easily tax the ability of an individual thinker, as anyone who has played the game knows. Task difficulty can be manipulated by altering the obscurity of the word to be guessed. In addition, task difficulty increases during the course of a single game as the amount of information that must be managed accumulates. We therefore predicted that groups would solve a higher proportion of games than individuals, and that the relative advantage of groups would increase for obscure words relative to simple words, and would increase during the later part of the game relative to the early part of the game.

All complex tasks—physical or cognitive—consist of a number of subtasks that must ultimately be isolated and studied in relation to each other for more complete understanding. We therefore conducted a second experiment in which part of what is required to solve the game of Twenty

Questions was presented in the format of a brainstorming experiment. Using the brainstorming format enabled us to relate our findings to a large body of previous research and to compare real groups with nominal groups in addition to single individuals.

Experiment 1: The Standard Game of Twenty Questions

*Methods.*The experiment was part of a semester-long course in which 36 undergraduate students first acted as participants and then helped to analyze the results while learning about decision making in a seminar format. Prior to the experiment, forty students from the Binghamton University Psychology Department's human subject pool were asked to write as many job titles as possible (defined as any activity that is done for payment and can be described in a single word) over a 40-minute period. These lists were merged and redundancies were removed to yield a master list of 442 job titles. The number of lists upon which a given job title appeared served as an index of its availability for recall in our subject population. For example, "doctor" appeared on all 40 lists while "bricklayer" appeared on only one list, even though "bricklayer" can easily be recognized as a job title once it is recalled. We deliberately chose the category of job titles for our experiment because it was familiar but presumably did not exist as an already organized category in the minds of our subjects, in contrast to a category such as mammals.

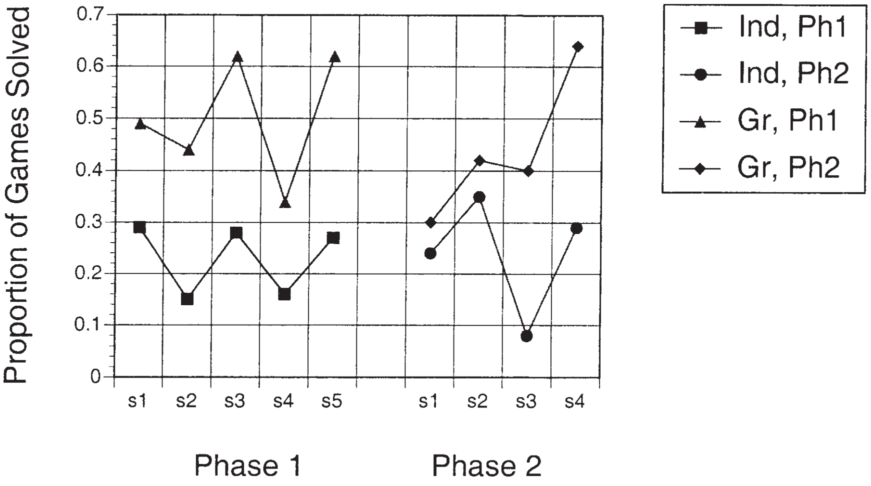
The experiment was divided into two phases. During Phase 1, half the participants were randomly assigned to same-sex groups of three people while the other half functioned as individuals. The groups and individuals played the game of Twenty Questions for five 1-hour sessions in which the job titles to be guessed were drawn randomly from the master list. They were told that the object was to guess the word using no more than 20 questions and without regard to time. Games that were in progress at the end of a session were discarded from the analysis. The games were conducted in two rooms similar to each other and with a minimum of objects present, to avoid cueing effects. Participants were instructed not to talk about the experiment with anyone outside their groups. Individuals and groups were read the same set of instructions at the beginning of each session, and groups were allowed to conduct their interactions as they saw fit. Writing was prohibited during the games. Phase 2 of the experiment consisted of four 1-hour sessions in which the Phase 1 groups were split into individuals and the Phase 1 individuals were formed into same-sex groups of three individuals. The experiment was conducted during a period of 5 weeks with two sessions per week.

*Analysis.*Our main dependent variable was the proportion of games solved by an individual or group during a single session. This unit of analysis avoids certain statistical biases by providing a single score per individual (or group) per session, regardless of how many games were played in a session. In addition to the final outcome of winning or losing, performance during the course of a single game was analyzed in the following manner. Consider an actual game in which the job title to be guessed is "bricklayer" and the first question asked is, "Does the job require a college degree?" This question can be answered not only for the job title of bricklayer but for all job titles on the master list, allowing the fraction of job titles excluded by the question to be determined. The second question—such as, "Is the job performed outdoors?"—can similarly be answered for all the job titles on the master list that were not excluded by the first question. In this fashion, a single game can be represented as a decay curve for the number of job titles on the master list remaining for consideration. The decay curve is steep for a well-played game and shallow for a poorly played game, regardless of whether the job title is actually guessed by the end of the game. This analysis required the construction of a large file in which all the questions asked by either individuals or groups for a number of games (more than 800 questions) were answered for all 442 job titles on the master list. This job was accomplished with the help of the students in the course after the experiment was over, and the decay curves for individual games were calculated with a computer program written for the purpose.

*Results.*Figure 1 shows the mean proportion of games solved by groups and individuals for both phases of the experiment. Analysis of variance (2 conditions x 2 phases, with sessions nested within phases) detected a main effect of group vs. individual (ANOVA, $F_{1, 212} = 16.56$, $p < 0.001$) but no significant main effect for Phase 1 vs. Phase 2 ($p = 0.94$), and no significant interaction ($p = 0.73$). There was also no difference between sessions within a phase ($p = 0.69$). Thus, groups performed roughly twice as well as individuals, but neither condition increased in performance over the course of the experiment. In addition, the experience of playing as a group during Phase 1 did not increase the performance of group members when they played as individuals in Phase 2. The advantages of thinking as a group evidently required actually being in a group.

Because there were no significant differences between phases or sessions, they were combined for subsequent analysis. Figure 2 ranks the performance of the twelve groups, from best to worst, along with performance of the best group member and the average performance of the three group members when playing as individuals. There is not even a

Figure 1. Mean proportion of games solved by individuals (Ind) and three-person groups (Gr) as a function of session number and phase. The members of groups during Phase 1 played as individuals during Phase 2 and vice versa. Phase 1 consisted of five 1-hour sessions (s) and Phase 2 consisted of four 1-hour sessions.



hint of a correlation between group and individual performance ($n = 12$, $r^2 = 0.014$, $p = 0.63$ for group vs. average member; $n = 12$, $r^2 = 0.020$, $p = 0.66$ for group vs. best member). Some of the best groups were composed of members who were mediocre as individuals and vice versa.

Figure 3 shows the mean decay curves for the first game of Phase 1, Session 1 and the first game of Phase 1, Session 5. The curves for individuals and groups are similar for the first five questions of the game but then diverge, with the groups exhibiting steeper curves than the individuals. There was a significant difference in number of job titles remaining prior to questions 6–12 and 18–20 (Kruskal-Wallace one-way ANOVA, $0.004 < p < 0.032$). Note that differences necessarily became smaller toward the end of the game owing to a “floor effect” in which the number of remaining job titles converges upon zero. Our interpretation of Figure 3 is that a certain degree of task difficulty is required for groups to outperform individuals. By its nature, the game of Twenty Questions becomes more difficult with every question because an accumulating amount of information must be remembered and evaluated to intelligently parse the remaining possibilities in framing the next question. Individuals were evidently as good as groups at asking the first few questions but then began to falter under the weight of the accumulating information load.

Figure 2. A comparison of group performance with the performance of the members playing as individuals. The groups are ranked according to group performance for convenience. There is no correlation between group performance (triangles) and the individual performance of either the average member (squares) or the best member (circles) of that group.

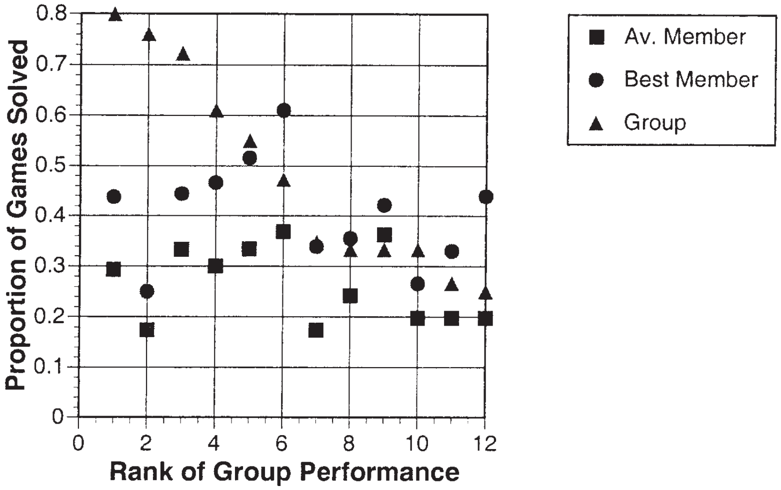
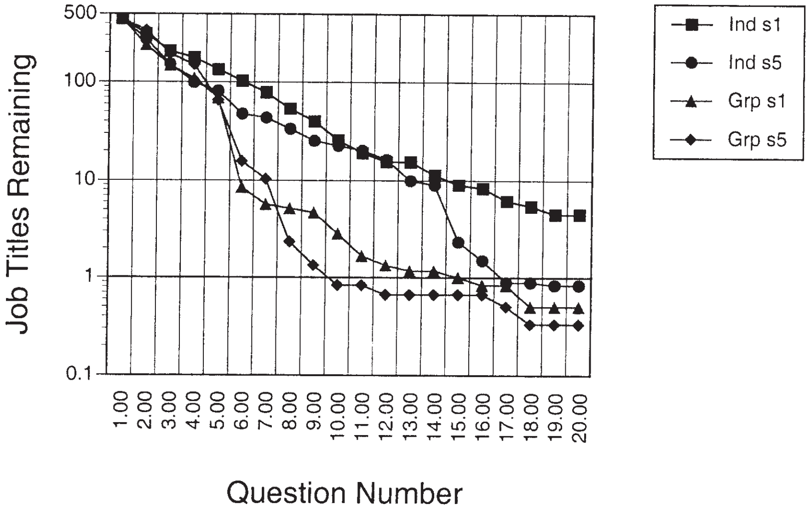


Figure 3. Mean number of job titles remaining as a function of the number of questions asked during a game of Twenty Questions. Squares and circles represent individual performance during Phase 1, Session 1 and Phase 1, Session 5, respectively. Triangles and diamonds represent group performance during Phase 1, Session 1 and Phase 1, Session 5, respectively.



The obscurity of the job titles to be guessed affords another way to measure task complexity. In principle it would have been possible to make obscurity a third factor of the experiment, along with individual vs. group and Phase 1 vs. Phase 2, but sample sizes did not permit a three-factor experiment. Instead, job titles were drawn at random from the master list for all games. To analyze the effect of obscurity, we split the job titles into two groups, those that appeared on 11–40 of the lists used to create the master list (relatively easy to recall) and those that appeared on 1–10 of the lists used to create the master list (relatively hard to recall). Figure 4 shows that groups surpass individuals for both easy and hard games. Groups and individuals both solve a lower proportion of hard games than easy games, but the performance decrease was 55% for individuals and only 32% for groups (all relevant pair-wise comparisons in Figure 4 are statistically significant: $\chi^2 = 4.32$–19.30, $p = 0.04$–0.001). This analysis does not permit us to attach statistical significance to the difference in the performance decrease (in contrast to an analysis of variance if obscurity had been included as a third factor, in which case the difference would appear as an interaction effect). Nevertheless, it qualitatively supports the conclusion that the advantages of cognitive cooperation increase with task difficulty.

Figure 4. Performance of individuals and groups for job titles that could be recalled with ease ("easy") or with difficulty ("hard"). Groups surpassed individuals and experienced a performance decrease of only 32% for hard words, compared with a 55% performance decrease for individuals.

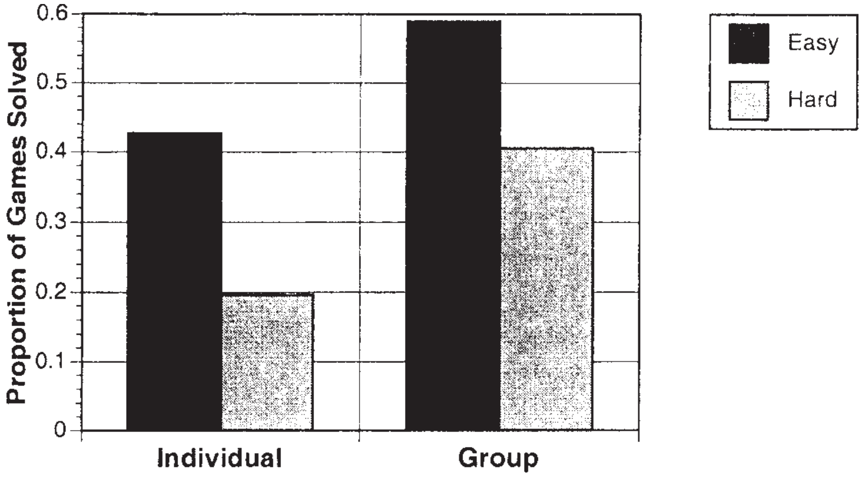
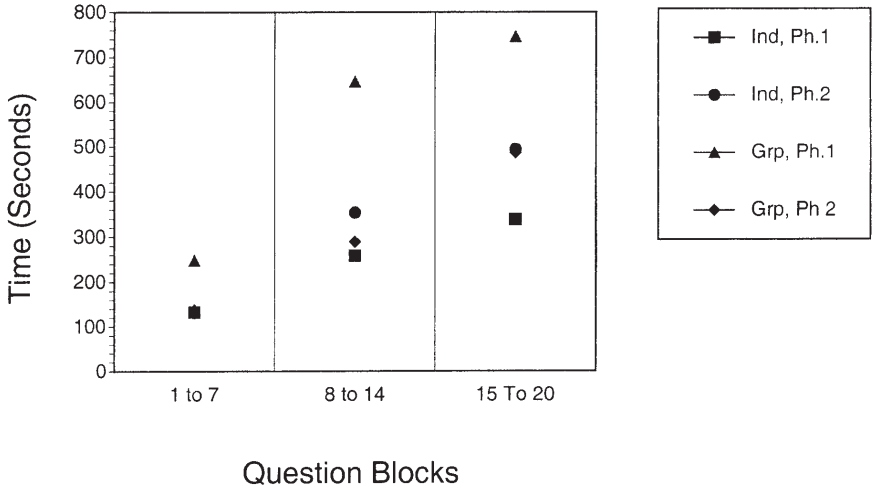


Figure 5. Mean time required to ask a question for individuals and groups during phase 1 and phase 2 of the experiment. The amount of time required to ask a question increased during the course of the game for both individuals and groups. Groups required more time than individuals for phase 1 but not for phase 2.



Individuals may not think *better* than groups, but it seems likely that they usually think *faster* than groups. Surprisingly, this reasonable expectation was confirmed for Phase 1 of our experiment but not for Phase 2, as shown in Figure 5. The difference in the amount of time spent on each question disappeared (perhaps because the participants were growing tired of playing the game of Twenty Questions) while the performance difference between groups and individuals remained.

Experiment 2: Partial Twenty Questions Game in a Brainstorming Format

We originally chose the game of Twenty Questions because we wanted to present individuals and groups with a challenging cognitive task. Not only did groups outperform individuals overall, but the details of their performance, as revealed in Figures 3 and 4, pointed to task difficulty as a critical variable for demonstrating the advantages of thinking in groups. However, the very complexity and multifaceted nature of the game made it difficult to identify the exact mechanisms that enhanced group performance. In addition, the most stringent test of group cognition is to compare the performance of real groups not with single individuals but with the same number of individuals thinking alone (nominal groups), as in brainstorming experiments. The game of Twenty Questions does not lend itself to the formation of nominal groups.

We therefore conducted another experiment to address these issues. Real groups and nominal groups composed of two people were asked to perform one of two tasks: (1) to think of as many job titles as possible, or (2) to think of as many job titles as possible that satisfy the criteria of a partially completed game of Twenty Questions in which seven questions had already been asked and answered. The second task is clearly more challenging than the first, although simpler than a complete game of Twenty Questions. We predicted that the increase in task complexity would enhance the performance of real groups relative to nominal groups. We also included a condition in which the task was performed by friends rather than strangers, to see if familiarity of group members had an effect on task performance.

*Methods.*One hundred eighty students (108 females, 72 males) from psychology classes at Binghamton University participated as part of their course requirements. Students were requested to sign up with a friend of the same gender if possible. The ideal experimental design would have included only pairs of friends, who would be kept together or split to form individuals and pairs of strangers. Such a design was not possible because an insufficient number of pairs of friends signed up for the experiment. We therefore paired friends with each other for the "friends" condition and used participants who signed up alone for the other conditions. This design leaves open the possibility that participants in the "friends" condition were drawn from a different sample population, but it does not affect the comparison of nominal groups with real groups of strangers. Participants who signed up as individuals were randomly assigned to nominal or real same-sex groups.

The experiment had three participant conditions (pairs of friends, pairs of strangers, and nominal pairs) and two levels of task difficulty (all jobs [easy] and partial game [hard]). Each individual participated in only one condition. Table 1 shows the questions and answers given to participants in the partial game condition, which were obtained from an actual game played during the previous Twenty Questions experiment. After signing the informed consent form, participants were given typed copies of the instructions, they followed along as the experimenter read the instructions aloud, and they were then asked if they had any questions. Next, they were provided 45 minutes to generate as long a list of single-word job titles as possible, either without constraints or with the constraints provided by the questions and answers of the partial game, which were provided on a written sheet. The experiment was conducted in four iden-

Table 1. A Partial Game of Twenty Questions Used for the Partial-Game Condition of the Brainstorming Study (Experiment 2). Real and nominal groups were asked to generate a list of job titles that satisfy these criteria.

1. Is it an office job?	No
2. Does the job involve contact with others?	Yes
3. Does the job require a college degree?	No
4. Does the job involve working with all ages?	Yes
5. Does the job involve working in a hospital?	No
6. Is it a government job?	No
7. Does the job involve working with food?	No

tical cubicles; trials on real and nominal groups were run in parallel whenever possible. After a sample size of approximately 10 pairs had accumulated for each of the six conditions, a preliminary analysis was conducted, prompting us to discontinue the "friends" and "all jobs" conditions and to double the sample size to approximately 20 pairs for the remaining partial-game conditions (real groups of strangers vs. nominal groups; see Methods). During analysis it became evident that four participants clearly misunderstood the instructions (e.g., by writing a list of job titles that satisfied the first question, followed by a list of job titles that satisfied the second question, and so on) and their lists were removed from the analysis.

That the "partial game" task is more difficult than the "all jobs" task may seem self-evident and is supported by comparing the lists generated for the two task conditions, as described in more detail below. Nevertheless, it seemed desirable to have at least some participants perform both tasks and subjectively rate their relative difficulty. Accordingly, five additional participants (beyond the 180 within the core part of the experiment) performed both tasks in random order for 10-minute periods and were asked to rate their relative difficulty (e.g., a rating of .5 indicates that the partial game was thought to be half as difficult as the full list, a rating of 2 indicates that the partial game was thought to be twice as difficult as the full list, and so on). These participants typed the job titles into a computer that was programmed to record the time of entry, enabling the rate of recall to be measured on a second-by-second basis along with the final number at the end of the tasks.

*Results.*In the first analysis of the data, real groups of strangers, real groups of friends, and nominal groups were found to perform equally well on the relatively simple all-jobs task ($F_{2,31} = 1.17$, $p = 0.32$; see Table 2). Performance on the more difficult, partial-game task was ranked in

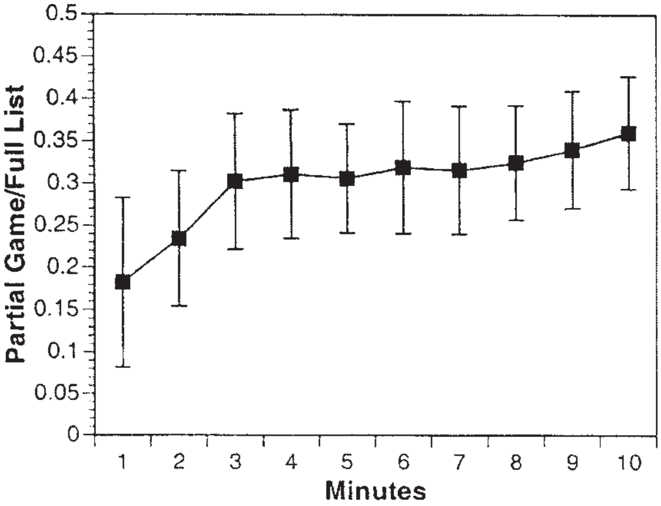
Table 2. Mean Number of Job Titles and Standard Errors Listed by Pairs of Strangers, Friends, and Nominal Groups of Two. The full-list condition involved guessing any one-word job title. The partial-game condition involved guessing one-word job titles that satisfied the conditions of the questions and answers shown in Table 1.

	First 63 Pairs						*Last 25 Pairs*		
	Full List			*Partial Game I*			*Partial Game II*		
	n	*mean*	*s.e.*	*n*	*mean*	*s.e.*	*n*	*mean*	*s.e.*
Strangers	11	187.0	13.3	10	96.7	16.6	11	93.1	13.74
Friends	11	177.4	10.2	9	77.8	11.9	—	—	—
Nominal	12	161.8	11.9	10	60.6	6.4	14	61.1	5.40

the order real groups of strangers > real groups of friends > nominal groups, although the differences did not reach significance at the 0.05 level ($F_{2,26}$ = 2.13, p = 0.14; restricting the comparison to real groups of strangers vs. nominal groups, $F_{1,18}$ = 4.02, p = 0.06). For both tasks, groups of friends did not perform better than groups of strangers. These results prompted us to increase the sample size for the most important comparison, that between real groups of strangers and nominal groups for the difficult task. The additional data alone demonstrated a significant performance advantage for real groups over nominal groups ($F_{1,23}$ = 5.72, p = 0.025). When the data for all participants were pooled, real groups of strangers had approximately a 50% performance advantage over nominal groups on the partial game task, a difference that is highly significant (with means of 94.8 vs. 60.8 items recalled; $F_{1,43}$ =10.20, p = 0.003).

All types of groups (real and nominal) listed half or less as many job titles for the partial-game task than the all-jobs task. By itself this is not surprising because fewer job titles qualified for the partial-game task. However, the lists from individuals who constituted nominal groups were actually less redundant for the partial-game task than for the full-list task (17.4% vs. 22.7%, respectively; $F_{1,34}$ = 11.48, p = 0.002). In other words, participants were far from exhausting the set of possible job titles for either task, but the process of recalling job titles for the partial-game task was slower and presumably more difficult. This interpretation was supported by the five additional participants who performed both tasks and whose rate of recall was measured throughout their 10-minute test period (Figure 6). The ratio of the number of words recalled for the partial-game task divided by the number of words recalled for the all-jobs task provides an index of the relative difficulty of the two tasks as a function

Figure 6. Ratio of the number of job titles recalled per minute for the partial game condition divided by the number of job titles recalled per minute for the all-jobs condition, providing an index of relative task difficulty. The ratio is lowest at the beginning of the recall period.



of time. The ratio is lowest at the beginning of the period, as participants struggled to assimilate the constraints imposed by the seven questions, and then rose to a plateau of approximately 0.30–0.35 after three minutes. The five participants subjectively rated the partial-game task as 2–5 times more difficult than the all-jobs task (mean = 2.6, s.e. = 0.6).

DISCUSSION

Cooperation is most beneficial for tasks that exceed the capacity of individuals. This statement might seem so obvious that it doesn't need to be made. However, the "unsettled state" of the literature that we described in the introduction requires a back-to-basics approach to cognitive cooperation in humans. We tested the most basic prediction for the game of Twenty Questions and found exactly what one might expect: Not only did groups outperform individuals overall, but their relative advantage increased with task difficulty, both within single games and between games that differed in terms of the obscurity of the target word.

In addition to this basic result, some of our more detailed results are anything but obvious: It is reasonable to expect individuals and groups to improve their performance during the course of playing many games, but no learning effect was observed. Similarly, it is reasonable to expect that the experience of playing in a group during Phase 1 would increase individual performance during Phase 2, but this was not observed. Finally, it is reasonable to expect that whatever makes an individual good at playing the game alone would also contribute to group performance, but this was not observed. These results can be summarized as follows: performance is not based on learning (at least over the short term), the advantages of groups require being in a group, and the mechanisms that contribute to group performance are not just a sum of the mechanisms that contribute to individual performance.

Brainstorming is one task in which traditional psychologists have looked very hard for cognitive cooperation and generally failed to find it. However, the fact that task difficulty has not been recognized and manipulated as an independent variable in brainstorming experiments shows how much a back-to-basics approach is needed. We manipulated task difficulty in a brainstorming format and found exactly what one would expect: Real groups outperformed nominal groups when the cognitive task was made sufficiently difficult. The fact that the friendship of group members did not have an effect on group performance will be discussed in more detail below.

We do not wish to imply that task difficulty is a single variable. There are many kinds of tasks that can be easy or difficult in different ways. Some very difficult tasks are still best performed by individuals (e.g., playing the violin), while some tasks that require groups are easy in the sense that they require minimal coordination (e.g., lifting a table), even though they are difficult in the sense of exceeding the weight or dimensional lifting capacity of a single person. Every physical task must be examined on its own terms for the potential costs and benefits of cooperation, and the same is true of cognitive tasks. Our experiments merely demonstrate the performance advantage of groups for the game of Twenty Questions and the effects of increasing task difficulty in a crude and poorly understood sense. Much more work will be required to understand the cognitive operations that comprise the game in detail and why they are improved by *co*operation.

It is clear—if only in retrospect—that cognitive cooperation should receive as much attention as physical cooperation in the study of our own species, no less than that of the social insects. For the rest of this paper we will consolidate our back-to-basics approach by addressing two questions: How can cognitive cooperation be incorporated into the field of evolutionary psychology? Why does the traditional psychological literature often give the impression that human groups function poorly as cognitive units?

Cognitive Cooperation and Evolutionary Psychology

One hallmark of evolutionary psychology is to study the mind as a collection of adaptations to the problems of survival and reproduction in ancestral environments. Although the "environment of evolutionary adaptedness" (EEA) is difficult to reconstruct, it is crudely approximated by modern hunter-gatherer societies. Timmel (2001) searched the electronic version of the Human Relations Area File (HRAF) using the key word "decision" to see how often human decision making takes place in a solitary vs. group context in hunter-gatherer and other non-technological societies. The answer is that decision-making almost invariably takes place in a group context, at least according to these ethnographic sources. The main exception is warfare, which sometimes places a premium on the speed of the decision. Similarly, Boehm (1996) searched the ethnographic literature for examples in which a group was faced with an important decision and an anthropologist was present to describe the decision-making process in detail. One example involved a natural disaster and two others involved decisions about going to war. In all three cases, the decisions were made by a lengthy and well-orchestrated group-level process. We do not mean to imply that all decision-making took place in groups in the EEA, much less that they were always made cooperatively. Clearly, individuals do make decisions by themselves, and those decisions are less likely to be noticed by ethnographers than decisions made by groups. Nevertheless, it is safe to say that decision-making and other forms of cognition have taken place largely in a group setting throughout our evolutionary history. Put simply, when our ancestors were thinking, they were usually socializing at the same time. This is an important precondition for the evolution of cognitive cooperation.

A second hallmark of evolutionary psychology is its reliance on theoretical models to determine what adaptations are likely to evolve in a given environment. It might seem that cognitive cooperation is more likely to evolve in social insect groups than human groups because genetic relatedness is usually higher in the former. However, genetic relatedness is only one of many factors that influences the evolution of cooperation. Other factors include the costs and benefits of cooperation, probabilities of repeated interactions, conditional behaviors such as the "tit-for-tat" strategy of game theory, social control mechanisms, and so on (Dugatkin 1997; Sober and Wilson 1998). When all of these factors are considered, the evolution of cognitive cooperation in humans becomes extremely likely theoretically. Cognitive cooperation often produces substantial gains for everyone at minimal individual cost. For example, in a group that must stay together, everyone gains from making a wise decision about where to move; the cost of the mental operations required to make the decision are energetically trivial and in any case are shared among the group members. This kind of low cost/high gain cooperation can evolve even in groups of unrelated individuals. In general, the cost/benefit ratio for cognitive cooperation might often be very low because the mental operations required for cooperation can be trivial in terms of time, energy, and risk, while the beneficial impact on behavior can be substantial.

We do not mean to imply that cognitive cooperation is invariably adaptive in human groups. Not only are many tasks performed best by individuals, as we have already stressed, but various kinds of free-riding and exploitation are possible, especially when there is a conflict of interest about the behavioral outcomes of cognition. Furthermore, even when cognitive cooperation is adaptive, the optimal group size might often be smaller than the actual size of the group. Just as some tasks are best performed by a single individual and a second would only get in the way, other tasks might best be performed by three individuals and a fourth would only get in the way. For all of these reasons, theoretically we should expect cognitive cooperation in humans to be richly context-sensitive and protected by social control mechanisms to eliminate the potential for exploitation. For example, pairs of friends did not perform better than pairs of strangers in our brainstorming experiment. However, the task was such that even strangers had everything to gain and nothing to lose by cooperating with each other (assuming that they were motivated to perform well at the task). If the task had been changed to provide an opportunity for exploitation, we predict that individuals might “shut down” in the presence of a stranger and “open up” in the presence of a friend, making friendship an important variable in cognitive cooperation. In general, context sensitivity provides many opportunities to formulate and test specific hypotheses about cognitive cooperation in humans based on theoretical models.

A third hallmark of evolutionary psychology is its emphasis on sophisticated, special-purpose adaptations that operate beneath conscious awareness (Barkow et al. 1992). Cognitive abilities such as vision and memory are “simple,” “effortless,” and “natural” in the sense that we do them without conscious effort and without the need for instruction, but the cognitive mechanisms that make these abilities so effortless are highly complex. We suggest that these same insights might apply to cognitive cooperation. Individuals playing the game of Twenty Questions by themselves frequently became exasperated and even downright stressed when they were “stumped,” “ran out of ideas,” and so on. They clearly had depleted their own cognitive resources in some sense and seemingly were yearning for external input. Some even described the experience of playing by themselves as "agonizing." In contrast, the members of three-person groups simply seemed to "click" into action, erupting into animated conversation, often with much gaiety and laughter. Individuals would impulsively barge into the conversation when they felt they had something to contribute, praise other members for ideas deemed valuable, and groan loudly when a promising idea proved to be wrong. Guessing the right word sometimes even resulted in cheers and the slapping of hands, as when a sports team scores a goal. These social dynamics are so familiar to us that we take them for granted, like sight and memory. However, the mechanisms that make them so effortless might be highly complex, as evolutionary psychologists already appreciate for individual-level cognitive adaptations (see Wilson et al. 2000 for a similar analysis of gossip as a group-level cognitive adaptation).

To summarize, there is every reason for cognitive cooperation to occupy center stage in human evolutionary psychology and to be integrated with the insights that are already being applied to the study of individual cognition.

Cognitive Cooperation and Traditional Psychology

The traditional psychological literature on thinking in groups has been reviewed from an evolutionary perspective by Wilson (1997) and consists of dozens of research programs on different cognitive tasks. These programs are often highly sophisticated and informative within their own domains but poorly integrated with each other, and some of the most basic questions that need to be asked from an evolutionary perspective often are not addressed. Evolutionary psychologists will be unsurprised to learn this because it is typical for nearly every major subject in psychology. Here we will summarize a few points from Wilson (1997) to show that the traditional psychological literature offers much support for the concept of cognitive cooperation, even though appearances are often to the contrary.

In the first place, some research programs do ask the right questions from an evolutionary perspective and impressively show the advantages of thinking in groups, especially for difficult tasks. Two outstanding examples are the work of Edwin Hutchins (1995) on cognition in naturalistic situations and Daniel Wegner on what he calls "transactive memory" (1986; Wegner et al. 1991). Hutchins uses navigation in both traditional and modern societies as a naturalistic setting for studying cognition. When a modern ship is a comfortable distance from shore, the task of sighting landmarks and marking the position of the ship on a chart can be accomplished by a single person. However, when the ship approaches the shore and its position needs to be determined at more frequent intervals and with less tolerance for error, this task is accomplished by a six-person team whose interactions have been refined over decades and even centuries of nautical history. Hutchins minutely analyzes these interactions in cognitive terms, much as social insect biologists examine the social cognition of bees and ants. Hutchins is one of the few to emphasize the transition from individual cognition to group cognition *as the task becomes more difficult,* which also forms the basis for our research. Wegner invites us to imagine how a computer engineer would connect a number of microcomputers, each with limited memory capacity, to form an integrated network with a larger memory capacity. That is how we should think about human memory according to Wegner, who supports his claim with a number of convincing laboratory studies. Wegner (1986) also reviews the history of thinking on group cognition in the social sciences, which should be read by everyone who wishes to integrate the psychological and evolutionary literatures for this subject.

Janis (1972, 1982) also studied group decision making in naturalistic situations, but he came to the opposite conclusion as Hutchins when he coined the word *groupthink:* "I use the term 'groupthink' as a quick and easy way to refer to a mode of thinking that people engage in when they are deeply involved in a cohesive in-group, when the members' strivings for unanimity override their motivation to realistically appraise alternative courses of action. Groupthink refers to a deterioration of mental efficiency, reality testing and moral judgement that results from in-group pressures" (Janis 1972:9).

Janis based his assessment on a qualitative and retrospective analysis of foreign policy disasters, but his work stimulated a more rigorous literature based on both laboratory experiments and the quantitative analysis of historical events. There are, of course, examples of both good and bad decision-making groups in American foreign policy history. For example, Lyndon Johnson was an overbearing leader who greeted dissenting views from his advisors with the ominous statement "I'm afraid he's losing his effectiveness," forcing many to leave his inner circle and those who remained to withhold their opinions. In essence, Johnson reduced the size of his decision-making group to himself. In contrast, the Marshall plan was formulated in a period of three weeks by a leader (George Kennan) who deliberately encouraged discussion and disagreement among his advisors. These and other historical examples have been analyzed with methods that illustrate the best of traditional social science research. They show that the quality of a group decision depends very much on how the group is structured and that groups dominated by a single individual usually do not make the best decisions. By itself, this is an argument for cognitive cooperation, not against it. However, even if groups can be structured to make better decisions than individuals, it is important to ask whether human groups spontaneously adopt the right structure, as we would predict from an evolutionary perspective and as our own research seems to indicate. Janis disagreed; he thought that human groups spontaneously adopt structures that lead to maladaptive outcomes. Subsequent research has proven Janis wrong on this point. When groups are made more cohesive and tasks are made more salient in laboratory experiments and in naturalistic situations, they become better decision-making units, not worse. As Aldag and Fuller (1993:539) conclude, "On the basis of our review, it seems clear that there is little support for the full groupthink model. . . . Furthermore, the central variable of cohesiveness has not been found to play a consistent role. . . . This suggestion is diametrically opposed to Janis's (1982) view that high cohesiveness and an accompanying concurrence-seeking tendency that interferes with critical thinking are 'the central features of groupthink'."

It is easy to sympathize with the appeal of the groupthink concept, since so many efforts to think in groups do appear dysfunctional in modern life, as anyone who has attended committee meetings can attest. However, evolutionary theory does not predict that groups are invariably better than individuals as cognitive units. There are plenty of situations in which individuals function better than groups, small groups function better than large groups, groups structured one way function better than groups structured another way, or lack of interest in the task and bitter conflicts of interest about the outcome turn groups into slumber parties or battlefields rather than cooperative units. Part of studying cooperative cognition from an evolutionary perspective involves appreciating its richly-context-sensitive nature, which makes simple generalizations impossible but provides the basis for many specific predictions. When we contemplate the efficacy of groups as cognitive units, we should think not only of boring committee meetings but also parties, scientific conferences, urgent councils of war, and hushed conversations with intimates. As we have already described, the participants in our experiment found the experience of playing the game of Twenty Questions far more enjoyable in groups than as individuals.

Groupthink is an outstanding example of a literature in traditional psychology that superficially portrays groups as dysfunctional cognitive units but upon closer examination does nothing of the sort. The brainstorming literature provides another example. Brainstorming researchers frequently lament that brainstorming techniques remain popular in business and industry even though experiments do not provide evidence for their efficacy. The implication is that the experience of business and industry is subjective and poorly controlled, and therefore wrong, while the psychological experiments are objective and carefully controlled, and therefore right. We agree that subjective experience is prone to a host of biases, but we also think that another difference separates brainstorming in the real world from brainstorming in laboratory experiments: the difficulty of the task. Until brainstorming researchers vary task difficulty as an independent variable, there is simply no way to relate the results of their experiments to the challenges of making decisions in the real world.

Another major problem concerns frames of comparison. From an evolutionary perspective, the most important comparison is an individual thinking alone vs. as a member of a socially interacting group. For brainstorming researchers, the most important comparison is between so-called real groups of interacting individuals and so-called nominal groups composed of the same number of individuals thinking by themselves. These two comparisons might seem the same, but a closer look reveals that “nominal” groups are socially interacting groups in their own right in which the experimenter functions as a member. First the experimenter takes the ideas compiled by members of the nominal group and merges them into a single list, throwing out the redundancies along the way. The fact that this list is longer (and no less creative) than the list compiled by real groups is used as evidence that nominal groups are superior to real groups. However, the real groups do not require the services of the experimenter to compile their list, and the time spent by the experimenter is not added to the time spent by the nominal groups. In addition, actually using the list to make a decision requires examining all the items on the list. The real groups can do this while compiling the list, whereas the nominal groups can’t even get started until after the experiment is over. When these factors are taken into account, it is not clear that a real group trying to make a real decision would want to emulate the structure of so-called nominal groups, even for the simple tasks employed in brainstorming experiments. Even if they did, they would be participating in a group-level cognitive process requiring social interactions between individuals in addition to neuronal interactions within individuals. The comparison of real vs. nominal groups is a comparison between two differently structured groups, not a comparison of individuals vs. groups. Individuals are so obviously inferior to either real or nominal groups in brainstorming experiments that the result seems unworthy of attention, yet it is exactly this “obvious” comparison that needs to be made from an evolutionary perspective and that reveals the advantages of individuals joining their minds with others to make better decisions than they can by themselves.

It is interesting to compare the brainstorming literature with computer algorithms that employ groups of cooperating agents to solve very complex problems (such as the travelling salesman problem), and which sometimes are inspired by cooperation in nonhuman species such as social insects (e.g., Bonabeau, Dorigo, and Theraulaz 2000) Typically these algorithms involve alternating between an autonomous phase in which the agents search the parameter space independently and an integrative phase in which the agents communicate to decide what to do next. The parallel processing that takes place during the autonomous phase is regarded as an advantage of group cognition that is unavailable to a single, sequentially processing unit. In brainstorming experiments, it is the members of nominal groups who enjoy the advantages of parallel processing while the members of real groups are forced into a sequential processing mode. Thus, what is properly regarded as an advantage of group cognition in one field of inquiry is associated with nominal groups (therefore an “individual” property) that is somehow unavailable to real groups in brainstorming experiments. This bizarre state of affairs can be avoided in the future by first comparing the performance of individuals who truly function alone with the performance of groups that themselves can differ in their social organization, of which so-called real and nominal groups count as two examples.

We do not mean to imply that the comparison between real and nominal groups is uninteresting. Having found that groups often surpass individuals as cognitive units, it becomes important to know exactly how group cognition works, what makes some social structures better than others, and especially the kinds of social structures that people adopt spontaneously. The brainstorming literature has been exemplary in dissecting the so-called performance deficits of real groups in comparison to nominal groups, and this surgical approach will become even more interesting when applied to tasks that allow real groups to surpass nominal groups, as in our second experiment, and especially for social structures that emulate what people adopt spontaneously. In a recent review of brainstorming, Brown and Paulus (2002:211) state: “It is clear that unstructured groups left to their own devices will not be very effective in developing creative ideas.” We predict that the very opposite conclusion will be reached when brainstorming and other cognitive tasks are studied from an evolutionary perspective.

We will provide one other example of how even the most clear-cut demonstration of cognitive cooperation in the psychological literature can be made to appear "individualistic" by altering the frame of comparison. Michaelsen, Watson, and Black (1989; see also Watson, Michaelsen, and Sharp 1991) taught college courses in which students were organized into learning groups that lasted the entire semester. The majority of class time was spent on group problem-solving tasks, including six objective and at least two essay exams that accounted for more than 50% of the course grade. Groups also met frequently outside of class to study and complete projects. This is one of the few research programs in which groups were presented with contextually relevant tasks and group dynamics had time to develop, as opposed to short-term laboratory experiments.

Exams were administered first to individuals and immediately afterward to groups (see Hill 1982 for a discussion of this and other experimental designs in group-decision research). In other words, after group members handed in their answer sheets, they were given an additional answer sheet for the same exam to fill out as a group. For a total of 222 groups from 25 courses taught over a 5-year period, the mean individual test score was 74.2, the mean score of the best individual in each group was 82.6, and the mean group score was 89.9. A total of 215 groups (97%) outperformed their best member, four groups (2%) tied their best member, and three groups (1%) scored lower than their best member.

These results provide overwhelming evidence that groups do not simply defer to their best member but scrutinize the questions on a case-by-case basis to decide which member is most likely to be correct. This is an impressive error-correcting capability that can take place only in groups. Nevertheless, the research was criticized by Tindale and Larson (1992a, 1992b), who claimed that groups must be able to answer a question right when every member answered it wrong to demonstrate a so-called "assembly bonus effect." Michaelsen, Watson, Schwartzkopf, and Black (1992) replied that their groups actually do display this effect, but our point is that the entire debate is peripheral from an evolutionary perspective. The ability to figure out which member is most likely to be correct by itself is an example of cognitive cooperation. The ability to figure out that no one was correct would be even more impressive, but is not necessary to show the basic advantage of thinking in groups.

These examples and others reviewed by Wilson (1997) show that the traditional psychological literature provides much evidence for cognitive cooperation, but only after it has been carefully reinterpreted from an evolutionary perspective.

We can summarize the results of our experiments, along with our review of the evolutionary psychology and traditional psychology literatures, as follows: The single human mind is adapted to function as a self-contained cognitive unit in some respects, but in others it is adapted to merge with other minds through social interactions to produce cooperative cognitive outcomes. This ability is surely a product of genetic evolution in addition to short-term learning and cultural processes. Cognitive cooperation needs to occupy center stage in evolutionary psychology, which in turn can provide a unifying conceptual framework for all research on group cognition.

We thank our 1996 decision-making class for participating in this research. We also thank the ecology, evolution, and behavior graduate group for helpful discussion.

David Sloan Wilson is an evolutionary biologist interested in a broad range of issues relevant to human behavior. He has published in psychology, anthropology, and philosophy journals in addition to his mainstream biological research. He is author of *Darwin's Cathedral: Evolution, Religion, and the Nature of Society* (University of Chicago Press, 2002) and co-author with philosopher Elliott Sober of *Unto Others: The Evolution and Psychology of Unselfish Behavior* (Harvard University Press, 1998).

John J. Timmel received his Ph.D. from Binghamton University in 2001.

Ralph R. Miller is Distinguished Professor of Psychology at Binghamton University. His research interests include information processing in animals, with an emphasis on elementary, evolutionarily derived, fundamentals of learning and memory that might be expected to generalize across species, including humans.

REFERENCES

Aldag, R. J., and S. R. Fuller
1993 Beyond Fiasco: A Reappraisal of the Groupthink Phenomenon and a New Model of Group Decision Processes. Psychological Bulletin 113:533–552.
Barkow, J. H., L. Cosmides, and J. Tooby, eds.
1992 *The Adapted Mind: Evolutionary Psychology and the Generation of Culture.* Oxford: Oxford University Press.
Boehm, C.
1996 Emergency Decisions, Cultural Selection Mechanics and Group Selection. *Current Anthropology* 37:763–793.
Bonabeau, E., M. Dorigo, and G. Theraulaz
2000 Inspiration for Optimization from Social Insect Behavior. *Nature* 406:39–42.
Brown, V. R., and P. B. Paulus
2002 Making Group Brainstorming More Effective: Recommendations from an Associative Memory Perspective. *Current Directions in Psychological Science* 11:208–212.
Camazine, S., J.-L. Deneubourg, N. R. Franks, J. Sneyd, G. Theraulaz, and E. Bonabeau
2001 *Self-organization in Biological Systems.* Princeton: Princeton University Press.
Cosmides, L., and J. Tooby
1992 Cognitive Adaptations for Social Exchange. In *The Adapted Mind,* J. Barkow, L. Cosmides, and J. Tooby, eds. Pp. 163–225. New York: Academic Press.

Dugatkin, L. A.
1997 *Cooperation among Animals.* Oxford: Oxford University Press.
Gigerenzer, G., and U. Hoffrage
1995 How to Improve Bayesian Reasoning without Instruction. *Psychological Review* 102:684–704.
Gigerenzer, G., P. M. Todd, and A. R. Group
1999 *Simple Heuristics That Make Us Smart.* Oxford: Oxford University Press.
Hill, G. W.
1982 Group versus Individual Performance: Are N+1 Heads Better Than One? Psychological Bulletin 91:517–539.
Hutchins, E.
1995 *Cognition in the Wild.* Cambridge, Massachusetts: MIT Press.
Janis, I. L.
1972 *Victims of Groupthink.* Boston: Houghton Mifflin.
1982 *Groupthink,* second ed. Boston: Houghton Mifflin.
Michaelsen, L. K., W. E. Watson, and R. H. Black
1989 A Realistic Test of Individual versus Group Consensus Decision Making. *Journal of Applied Psychology* 74:834–839.
Michaelsen, L. K., W. E. Watson, A. Schwartzkopf, and R. H. Black
1992 Group Decision Making: How You Frame the Question Determines What You Find. *Journal of Applied Psychology* 77:106–108.
Mullen, B., C. Johnson, and E. Salas
1991 Productivity Loss in Brainstorming Groups: A Meta-analytic Integration. *Basic and Applied Social Psychology* 12:3–24.
Osborne, A. F.
1957 *Applied Imagination.* New York: Scribners.
Seeley, T.
1995 *The Wisdom of the Hive.* Cambridge: Harvard University Press.
Seeley, T., and S. C. Buhrman
1999 Group Decision Making in Swarms of Honey Bees. *Behavioral Ecology and Sociobiology* 45:19–31.
Sober, E., and D. S. Wilson
1998 *Unto Others: The Evolution and Psychology of Unselfish Behavior.* Cambridge: Harvard University Press.
Stroebe, W., and M. Diehl
1994 Why Groups Are less Effective Than Their Members: On Productivity Losses in Idea-Generating Groups. *European Review of Social Psychology* 5:271–303.
Taylor, D. W., and W. I. Faust
1952 Twenty Questions: Efficiency in Problem Solving as a Function of Size of Group. *Journal of Experimental Psychology* 44:360–368.
Timmel, J. J.
2001 *Group Cognition from a Multilevel Evolutionary Perspective.* Ph.D. dissertation, Binghamton University.
Tindale, R. S., and J. R. J. Larson
1992a Assembly Bonus Effect or Typical Group Performance? A Comment on Michaelsen, Watson and Black (1989). *Journal of Applied Psychology* 77:102–105.
1992b It's Not How You Frame the Question, It's How You Interpret the Results. *Journal of Applied Psychology* 77:109–110.
Watson, W., L. K. Michaelsen, and W. Sharp
1991 Member Competence, Group Interaction, and Group Decision Making: A Longitudinal Study. *Journal of Applied Psychology* 76:803–809.

Wegner, D. M.
1986 Transactive Memory: A Contemporary Analysis of the Group Mind. In *Theories of Group Behavior,* B. Mullen and G. R. Goethals, eds. Pp. 185–208. New York: Springer-Verlag.
Wegner, D. M., R. Erber, and P. Raymond
1991 Transactive Memory in Close Relationships. *Journal of Personality and Social Psychology* 61:923–929.
Wilson, D. S.
1997 Incorporating Group Selection into the Adaptationist Program: A Case Study Involving Human Decision Making. In *Evolutionary Social Psychology,* J. Simpson and D. Kendricks, eds. Pp. 345–386. Mahwah, New Jersey: Erlbaum.
Wilson, D. S., C. Wilczynski, A. Wells, and L. Weiser
2000 Gossip and Other Aspects of Language as Group-Level Adaptations. In *Cognition and Evolution,* C. Heyes and L. Huber, eds. Pp. 347–365. Cambridge, Massachusetts: MIT Press.