Group Decision-Making on an Optimal Stopping Problem

Michael D. Lee
Department of Cognitive Sciences
University of California, Irvine

Michael J. Paradowski
Command and Control Division
Defence Science and Technology Organisation

We consider group decision-making on an optimal stopping problem, for which large and stable individual differences have previously been established. In the problem, people are presented with a sequence of five random numbers between 0 and 100, one at a time, and are required to choose the maximum of the sequence, without being allowed to return to earlier values in the sequence. We examine group decision-making on these problems in an experimental setting where group members are isolated from one another, and interact solely via networked computers. The group members register their initial accept or reject decision for each value in the sequence, and then providing a potentially revised decision having viewed the recommendations of the other group members. Group decisions are made according to one of three conditions, requiring either consensus to accept from all group members, a majority of accept decisions from the group, or the acceptance of an appointed group leader. We compare individual decision-making to group decision-making under these three conditions, and find that, under some conditions, groups often significantly outperform even their best members. Using a signal detection analysis we provide an account of how the group decision-making conditions differ from one another, and from individual decision-making. Key findings are that people do not often revise their decisions, but, in the consensus and leadership conditions, are more conservative in their initial decisions. This conservatism removes the individual bias towards choosing values too early in the sequence, allowing the groups to perform better than their individual members. In the majority condition, however, people continue to behave as they did individually, and the group shows the same bias in decision-making.

Introduction

Optimal Stopping Problems

Most human decision-making can be conceived as searching through a sequence of alternatives until a choice is made. Often the number of possible alternatives considered is relatively small, because there are limited options in the external task environment, or because of the need to make fast decisions in a competitive world. In some situations, it is also not possible to re-consider a previously rejected alternative. In dynamic environments, previous evaluations may no longer be accurate, or—think, for example, of mate selection—the earlier act of rejection may incur large costs that make reconsideration prohibitive.

A class of optimization problems, generically known as optimal stopping problems (see Ferguson, 1989, for a historical overview), have features that make them well-suited to studying human decision-making on limited sequences of alternatives. For this reason, these problems have received steady theoretical and empirical attention over a long period in cognitive psychology (e.g., Bearden, Murphy, & Rapoport, 2005; Corbin, Olson, & Abbondanza, 1975; Dudey & Todd, 2001; Kahan, Rapoport, & Jones, 1967; Lee, 2006; Seale & Rapoport, 1997, 2000; Rapoport & Tversky, 1970) and other fields, such as experimental economics (e.g., Cox & Oaxaca, 1992; Kogut, 1990; Zwick, Rapoport, Lo, & Muthukrishnan, 2003)

In this paper, we consider human performance—both as individuals, and in various group settings—on an optimal stopping problem where people are presented with a list of five randomly chosen numbers between 0 and 100. People are told there are five numbers in the list, and they were chosen randomly. Individuals or groups are then shown the numbers one at a time, and are instructed to choose the maximum, subject to the constraint that they must choose a number at the time it is presented, and that any choice below the maximum is incorrect.

Gilbert and Mosteller (1966) provide an integrated overview of mathematical results for optimal stopping problems. Most interestingly, they describe the optimal decision process, the adherence to which maximizes the probability of making the correct choice for any randomly generated problem. This optimal decision-making process is to choose the first value that is both the maximum value observed in the sequence thus far *and* exceeds a threshold level for its position in the sequence. Gilbert and Mosteller (1966, Tables 7 and 8) provide these optimal thresholds and the associated probabilities of making a correct decision.

As a concrete example, Figure 1 shows a five-point problem, with the circles representing successive values in the problem, and the solid line showing the optimal threshold for each of the five positions (since the last value is a forced choice, its threshold is effectively zero). In this example, the optimal choice is the third value presented, as it is the maximum value seen to that point in the sequence, and is above

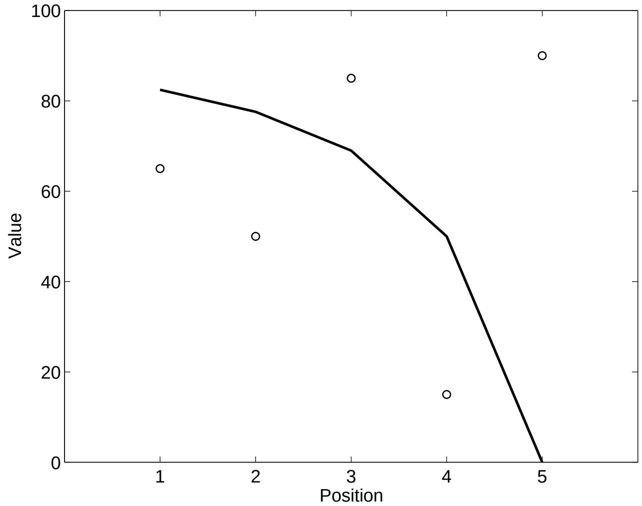

Figure 1. An example optimal stopping problem, showing the sequence of five values between 0 and 100, and the curve corresponding to the optimal decision process.

the threshold defined by the optimal curve. Note, however, that this choice is incorrect in the sense that it does not correspond to the maximum value in the sequence, which occurs in the fifth and final position. In this way, as argued previously by Lee (2006), optimal stopping problems naturally distinguishing between performance based on achieving optimal outcomes (i.e., choosing the final value), and performance based on following optimal decision processes (i.e., choosing the third value). Simon (1976) termed these different measures 'procedural' and 'substantive' rationality, respectively, and noted that procedural measures are inherently less noisy, because the optimal decision process can always be followed, even when optimal outcomes cannot always be achieved.

Individual Decision-Making

Most of the previous research examining human decision-making on optimal stopping problems has used versions of the problem that provide rank order information, rather than the values themselves (e.g., Dudey & Todd, 2001; Seale & Rapoport, 1997, 2000). These rank order problems, however, have a very different optimal decision rule, and so it is unclear to what extent their findings generalize to the current context. Kahan et al. (1967) did study human decision-making on a more similar partial-information task, where values rather than ranks are presented, but the distribution is not explicitly given to participants. These authors used problems of length 200, withifferent problems involving values drawn from either a positively skewed, negatively skewed, or a uniform distribution. No evidence was found for the different distributions affecting the decisions made. Corbin et al. (1975) considered human decision-making on problems like ours and, by systematically manipulating the values presented, found sequential and contextual dependencies within problems. Other empirical studies (e.g., Cox & Oaxaca, 1992; Kogut, 1990; Rapoport & Tversky, 1970; Zwick et al., 2003) have used very different experimental methodologies, such as requiring subjects to expend resources to consider additional alternatives, usually because they are interested in applications to economic decision-making.

The series of studies most directly relevant to the current one were conducted by Lee, O'Connor, and Welsh (2004), Lee (2006), and Campbell and Lee (in press). Lee et al. (2004) considered human performance on problems with lengths 10, 20 and 50, and evaluated three candidate models of the way people made decisions. They concluded that the best accounts were provided by 'threshold' models in which people choose by comparing the presented value to fixed thresholds. What Lee et al. (2004) observed, however, was that there seemed to be significant individual differences in the exact thresholds that people used. Some subjects behaved consistently with applying a single fixed threshold across the entire sequence. Effectively, these people chose the first number that exceeded a fixed value. Other subjects, however, behaved consistently with using thresholds that decreased as the sequence progressed, as with the optimal solution.

Lee (2006) examined the possibility of individual differences in more detail, observing that, over a total of 147 participants, each completing one of two different sets of 40 problems, there was evidence of individual differences, but no evidence of learning. In other words, the proportion of times the optimal solution process was followed differed between participants, but did not appear to change as the same participant answered additional problems. In addition, based on a model of the decision-making process, Lee (2006) was able to make inferences about the various thresholds used by people, and observed a wide variety of different types of solution processes being employed. Campbell and Lee (in press) provided additional evidence of the stability of these individual differences by testing a total of 75 participants on 120 problems of length five, under various feedback and financial incentive conditions, and observing no evidence of learning in any of the conditions.

Group Decision-Making

The finding of large and stable individual differences in decision-making raises a number of interesting questions about how groups will solve optimal stopping problems. Because people make different decisions as individuals, group decision-making must involve some sort of compromise across, or competition between, alternative answers. And, because people show few signs of learning or changing their decision-making on these problems over repeated trials, it is not obvious how such compromise or competition will be resolved.

A further attraction of studying group behavior on the optimal stopping problem is that has many desirable properties previously identified in the group decision-making literature. As Gigone and Hastie (1997) point out, most laboratory tasks involving group decision-making have required background knowledge, which is difficult to quantify. In contrast, the lack of background knowledge required to solve optimal stopping problems makes them amenable to quantitative analysis. In addition, an important question in the study of group decision-making is whether groups attenuate or exacerbate individual decison-making bias (see Kerr & Tindale, 2004, p. 634). To do this, as noted by Gigone and Hastie (1997), it is necessary to be able to collect repeated measures of individual and group decision-making. The optimal stopping task is also well suited to these demands. It is straightforward to generate and administer large numbers of essentially equivalent but new problems.

As far as we are aware, however, group decision-making on optimal stopping problems has never been considered experimentally (see the thorough experimental reviews in Gigone & Hastie, 1997; Hastie, 1986). The only previous empirical study to consider group effects of any form is that reported by Kahan et al. (1967). These authors compared the performance of individuals making optimal stopping decisions in an isolated setting with those still making decisions as individuals, but in a group setting under the condition that they had to remain in the experimental setting until the entire group had completed their problems. Not surprisingly, they found that in the group setting people chose to accept values earlier in the sequence than they did in isolation.

In this paper, we examine the decision-making of individuals, and groups of five people, completing five-point optimal stopping problems. We consider three within-group manipulations, involving consensus, majority and leadership-based decision-making for the group. Following the framework suggested by Gigone and Hastie (1997), we distinguish between 'individual' decisions made in isolation, 'member' decisions made at the beginning of a group process, 'revised member' decisions made after interaction with the other members of a group, and the final 'group' decision. We adopt a signal-detection theory approach to provide measures of both accuracy and bias on detailed decision-by-decision performance. We then use these measures to examine how decision making evolves in a group setting, how different group decision processes differ from one other, and how they differ from those of individual decision-makers.

Experiment

Participants

We tested seven groups of five participants, comprised of 13 male and 22 female participants, with an average age of 24.4 (SD=9.10) years. Participants were randomly assigned to groups, with gender and age distributions that broadly matched those of the entire sample.

Procedure

Individual Setting. Participants first completed a set of 20 problems working as individuals. For each problem participants were sequentially presented with numbers ranging from 0.00 to 100.00, and were instructed to choose the maximum value. It was emphasized that (a) the values were uniformly and randomly distributed between 0.00 and 100.00, (b) a value could only be chosen at the time it was presented, (c) the goal was to select the maximum value, with any selection below the maximum being completely incorrect, and (d) if no choice had been made when the last value was presented, they would be forced to choose this value. As each value was presented, its position in the sequence (e.g., the information that "this is the third number out of five") was shown, together with 'yes' and 'no' response buttons. No feedback was provided, and the order of the problems was randomized for each participant.

Group Setting. Participants then completed a total of 30 problems working as a member of a five-person group, with each person located remotely at a computer terminal and interacting only through the networked software that ran the experiment. For each successive number in each problem, this software showed the number, and its position in the sequence, to all members of the group, and asked for a member accept or reject decision. This decision was made by each group member in isolation, without knowledge of the decisions of the other members. Once all member decisions had been made, the software provided a graphical representation of the decisions to all group members. Each participant was then asked for a revised member accept or reject decision for the same number.

Over their experimental session, each group operated under three decision-making conditions we call 'consensus', 'majority' and 'leadership' conditions, and did ten problems in each condition. In the consensus condition, everybody in the group was required to make an accept decision at the member stage for that value to be chosen by the group as a whole. In the majority condition, three or more of the group had to accept the value for it to be chosen by the group. In the leadership condition, the one group member who was appointed leader made a decision at the member stage that became the group decision for that value. Leaders were assigned at random, and were changed, without re-selecting the same person, for each problem set. Whatever the condition, the accept or reject decision generated by each group for each value was treated in the same way as the individual decision-making setting. That is, groups continued to be presented with values in the problem sequence until one was selected, or the last value became a forced choice.

The basic group decision-making process is summarized in Figure 2. The five member of the group are shown, making decisions in relation to the presented value. A sample progression through member to revised member decisions is shown. From the revised member decisions, the group decision is determined by the consensus, majority or leadership condition rule. For this reason, in the leadership condition, a revised member decision was only required from the assigned leader. Each group did different randomly generated problems, and the order of the decision-making conditions was counterbalanced, to the maximum extent possible, across groups.

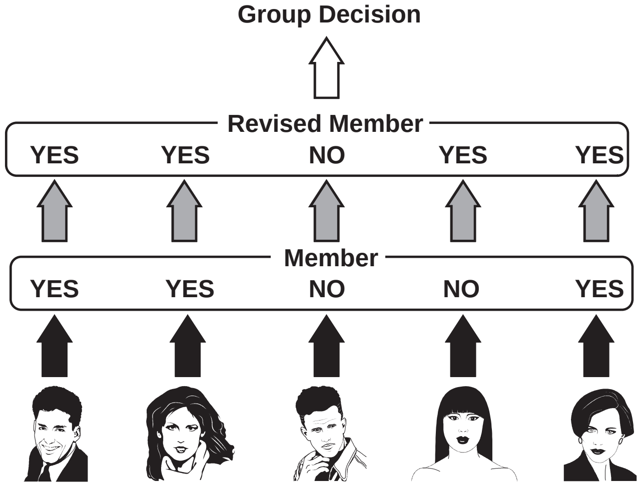

Figure 2. Overview of the basic experimental procedure for group decision-making, showing an example sequence of member and revised member decisions, from which the group decision is determined.

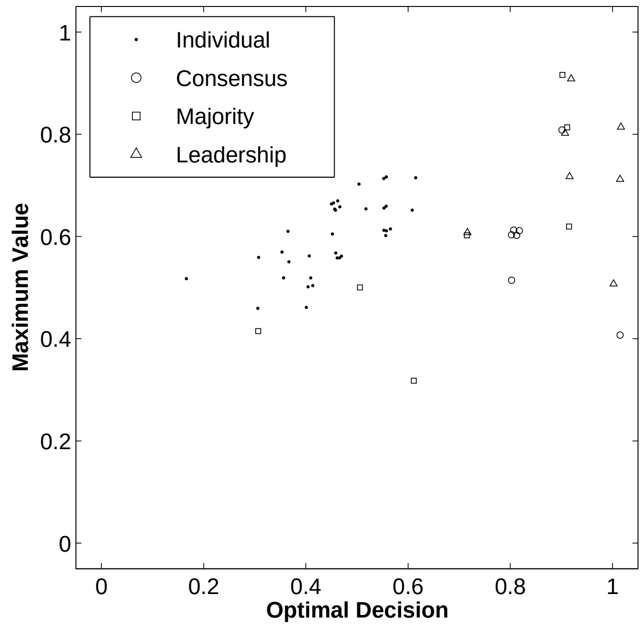

Figure 3. Proportion of optimal decisions against proportion of maximum values chosen, for individual participants, and for each decision-making condition. Each point has been slightly peturbed for visibility.

Basic Results

Figure 3 summarizes the accuracy of the decisions made by individuals and groups, both in terms of making optimal decisions, and in choosing the maximum value. This analysis makes clear a few basic conclusions. First, there are large differences in accuracy between individuals, and between groups using the same decision-making method. Any analysis of decision-making accordingly needs to accommodate individual differences. Secondly, there seem to be differences between the accuracy of groups and those of individuals. In particular, many consensus and leadership groups adhere perfectly to the optimal decision process, a feat no single individual achieved. Taken together, these observations suggest that there are difference between individual and group decision-making, and between different group decision-making conditions.

Signal Detection Analysis

We rely on a form of signal detection analysis (e.g., Green & Swets, 1966) to explore the differences in individual and group decision-making in more detail. The behavioral data we use are the individual, member, revised member, and group accept and reject decisions. As shown in Table 1, by comparing each accept and reject decisions to those required by the optimal decision process, we obtain counts of hits, false alarms, misses, and correct rejections. From these counts, it is straightforward to make inferences about the hit rate θ_h and false alarm rate θ_f of a decision-maker. Analyzing these data provides a much more detailed characterization of the decision-making process than simply considering the final decision.

Table 1
Signal detection table relating human decisions to the optimal decision process.

		Optimal Decision	
		accept	reject
Human	accept	hit	false alarm
Decision	reject	miss	correct rejection

Standard Signal Detection Theory

We assume the standard Gaussian equal-variance form of signal detection theory shown in Figure 4, because it allows the hit and false alarm rate information to be converted to useful and interpretable measures of decision-making performance. In signal detection theory, two stimulus distributions are proposed, representing signal and noise, separated by some distance d' that measures their discriminability. In general, the task of the decision-maker is to distinguish signal from noise stimuli. Signal detection theory assumes, in attemting to achieve this, that on each trial the decision-maker samples the stimulus from the signal or noise distribution as appropriate, and compares it to a referent criterion k. If the stimulus sample is greater than the criterion, the stimulus is identified as signal, otherwise it is identified as noise. The criterion k can be re-expressed in terms of a bias measure β, which is the ratio of the density of the signal to noise distributions at k, and in terms of the difference c between the k and the unbiased criterion value.

For our optimal stopping problem, the signal distribution

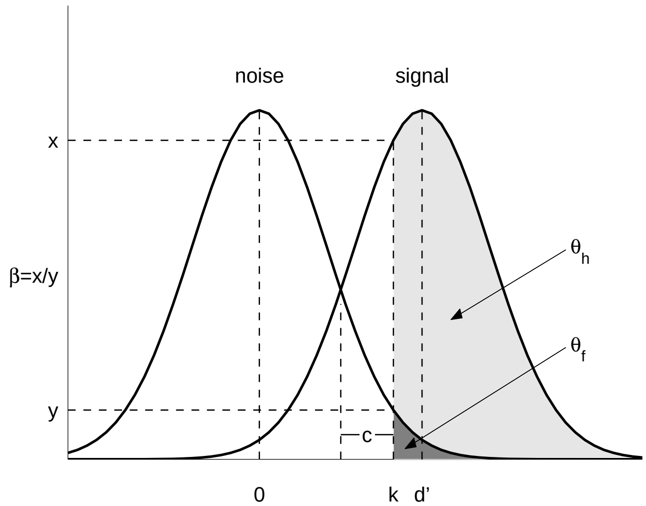

Figure 4. Standard signal detection theory framework.

corresponds to those values large enough for their position in the sequence to warrant an accept decision according to the optimal decision rule. The noise distribution corresponds to values that ought to be rejected. Discriminability now provides a direct measures of how well a decision-maker is performing in making accept and reject decisions, and so assesses how well they are adhering to the optimal decision process.

Because the signal distribution corresponds to those that should be accepted, the criterion k is not a fixed value against which the presented numerical values are compared. Rather, it is defined relative to the optimal decision process. It is also important to note that the unbiased criterion, and hence the definition of bias given by c, depends on the base rates for values that should be accepted and rejected. If, for example, there are four times as many values to be rejected as accepted, the optimal criterion is the one that gives a β value of four. We determined the actual base rates for our five-point problems, in each position, by simple Monte Carlo simulation.

Accordingly our measure of bias c is the difference between the actual criterion k and the unbiased criterion value, defined so as to take into account the position in the sequence. To the extent that a decision-maker does not follow the optimal decision process, the bias measure indicates whether they are tending to accept values that they should reject, corresponding to negative bias in our formulation, or reject values they should accept, corresponding to positive bias.

Extending Signal Detection Theory to Groups

While standard signal detection provides an account of the discriminability and bias of a single decision-maker (whether a single individual, or a single group), it does not provide any formal account of a collection of decision-makers. Given the large individual differences already noted, we want to compare the *distributions* of discrimability and bias for collections of individual, member, revised member and group decisions. To achieve this, we use a hierarchical Bayesian signal detection theory framework Rouder and Lu (2005). The technical details of our statistical methods are available as an on-line technical note from the first author's web page.

The hierarchical model extends basic signal detection theory by including an extra level of representation that describes how the discriminability and bias characteristics for a collection of decision-makers are distributed. Specifically, we assume these distributions are Gaussians, and that discriminability and bias are independent. Using standard statistical methods, we can then make inferences from the counts in Table 1 about the discriminability and bias of the decision-maker, but also about the mean and variance of the discriminabilities and biases of a collection of decision-makers. In turn, we can use standard statistical methods to test whether two collections of decision-makers are the same or different in their discriminability and bias distributions.

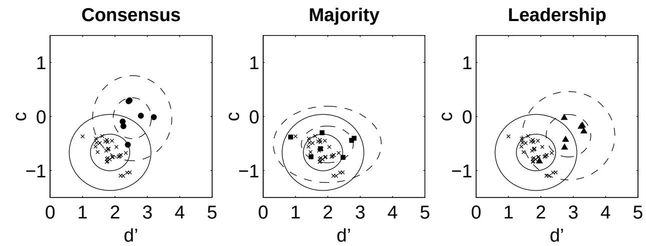

Figure 5. Individual and group behavior under each of the decision-making conditions. Within each panel, markers show the expected discriminability (d') and bias (c) derived from the decisions made by each individual and each group. Superimposed are the 50% and 95% contours for the expected Gaussians distribution over individuals and the groups.

Individual and Group Behavior

Figure 5 summarizes the results of applying the hierarchical signal detection model to the individual decisions and group decisions. The three panels correspond to the consensus, majority and leadership group decision-making conditions. Within each panel, crosses show the expected discriminability (d') and bias (c) derived from the decisions made by each of the 35 individuals, and circular, square or triangular makers show the expected discriminability and bias for each of the seven groups. Also shown are the 50% and 95% contours for the expected Gaussian distributions over the individuals and the groups.

Table 2 details the Bayes Factors (e.g., Kass & Raftery, 1995) that test whether the discriminability and bias distributions are the same or different in each case. The Bayes Factors are measured on the often-used logarithmic scale. On this scale, zero is the point of indifference: the point at which the data provide as much evidence for the distributions being the same as they do for the distributions being different. Positive values indicate evidence in favor of the distributions being the same, while negative values indicate evidence of a difference. Because the values themselves are simply the logarithm of a likelihood ratio, they are readily interpreted. We follow the suggested guide of Kass and Raftery (1995), where (absolute) values less than one are regarded as "not worth more than a bare mention", values between one and three are regarded as "positive", between three and five are regarded as "strong", and larger than five are regarded as "very strong". We are particularly interested in cases where individual and group decision-making differ, and so Table 2 highlight in bold those log Bayes Factors that are negative, with a magnitude greater than one.

Table 2 shows that the consensus and leadership groups have different levels of both discrimation and bias to individuals. With reference to Figure 5, it is clear that discriminability improves in both groups settings. It is also evident that a large negative bias for individuals is reduced to something close to an unbiased state in the consensus condition, and is also reduced, but to a lesser extent, in the leadership condition.

Individual Behavior in Groups

To consider the sequence of decisions each participant made—moving from their decisions as individuals to their their member to their revised member decisions in group settings—we use a 'within-subjects' version of the hierarchical signal detection analysis. This involves, instead of considering separate discriminability and bias measures for both member and revised member decisions, considering the change in discriminability $\Delta d'$ and change in bias Δc between these stages for each individual.

Figure 6 summarizes the results of applying the hierarchical signal detection model to the individual to member changes. Table 2 gives the Bayes Factors, which compare an account that assumes there is no change in discriminability and bias, with one that does allow for the change. As before, the Bayes Factors are measured on the log scale, and negative values indicate evidence for change. From these analyses, it is clear that in both the consensus and leadership decision-making conditions, but not in the majority condition, there is a change in discriminability and bias. In particular, the decisions people make as members show greater discriminability. It is also clear, with reference to Figure 5, that the increase in the value of the bias measure in the consensus and leadership condition has the effect of making the member decisions much closer to being unbiased than the individual decisions.

Figure 7 summarizes the results of applying the hierarchical signal detection model to the member to revised member changes, and Table 2 again gives the log Bayes Factors. These analyses suggest that there are no significant changes in either discriminability or bias as people move from making member to revised member decisions.

Analysis of Changes

The finding that there are no major changes in discriminability or bias in revising member decisions does mean that it is not worth examining those changes that do occur. Such

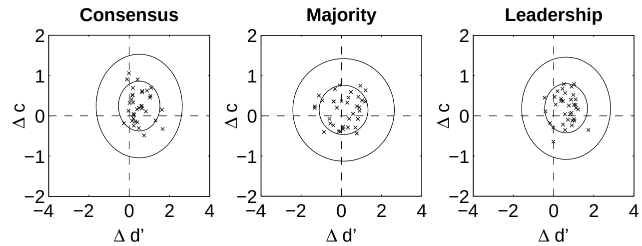

Figure 6. Changes from individual to member behavior under each of the decision-making conditions. Within each panel, markers show the expected change in discriminability ($\Delta d'$) and change in bias (Δc) derived for each participant moving from their individual to their member decision-making. Superimposed are the 50% and 95% contours for the expected Gaussian distribution over the collection of differences.

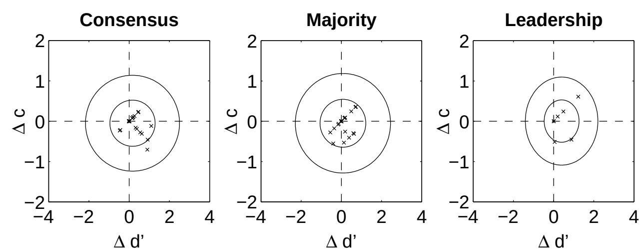

Figure 7. Changes from member to revised member behavior under each of the decision-making conditions. Within each panel, markers show the expected change in discriminability ($\Delta d'$) and change in bias (Δc) derived for each participant moving from their member to their revised member decision-making. Superimposed are the 50% and 95% contours for the expected Gaussian distribution over the collection of differences.

an analysis is presented in Figure 8, which shows the proportion of changes, relative to the total number of decisions in that condition, in each decision-making condition. These changes are shown according to whether they are 'good' changes (i.e., changes that changed a member decision not in accord with the optimal rule into a revised member decision that was in accord), or 'bad' changes (i.e., changes away from a member decision in accord with the optimal rule). These good and bad changes are shown further divided into those where the subject was 'encouraged' to change a member reject into a revised member accept decision, and those where the subject was 'discouraged' to change a member accept into a revised member reject decision.

Figure 8 shows that, under the consensus and majority condition, only about 15% of decisions were changed moving from the member to the revised member stage of the decision-making process. In the leadership condition, the leader changed their member decision about 20% of the time. These changes were much more often good changes than bad ones, especially in the leadership condition. The good changes were more often discouragements than encouragements, again especially in the leadership condition. And, finally, bad changes were almost exclusively encouragements.

Table 2
Log Bayes factors testing whether individual vs group, individual vs member, and member vs revised member decision-making have the same or different discriminability (d') and bias (c) characteristics. Positive values give evidence in favor of sameness; negative values give evidence in favor of differences. Negative values indicating substantial differences are highlighted in bold.

	Discriminability (d')			Bias (c)		
	Consensus	Majority	Leadership	Consensus	Majority	Leadership
individual vs group	**-7.10**	1.89	**-7.64**	**-4.97**	0.98	**-1.83**
individual vs member	**-3.66**	0.81	**-5.53**	**-4.72**	0.66	**-2.66**
member vs revised member	-0.24	1.82	0.60	0.66	0.37	0.61

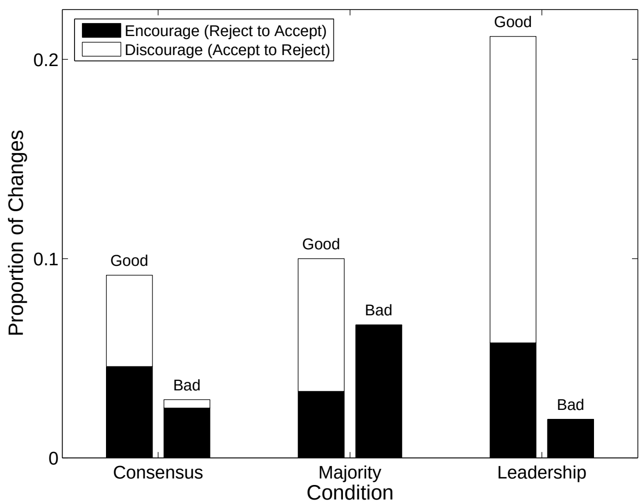

Figure 8. Analysis of changes in member to revised member decisions. The three panels show, top to bottom, the consensus, majority, and leadership conditions. Each panel shows the number of 'good' and 'bad' changes for five, ten and twenty point problems. These counts are further divided into how many changes were 'encouragements' versus 'discouragements'.

Individual Learning

Our final analysis examines the possibility that individuals learned while completing their 20 problems. While previous results strongly suggest there will be no learning, it is an important check, because otherwise the comparison of group and individual decision-making would be confounded with practise effects. Figure 9 shows the results of a within-participants hierarchical signal detection analysis of the change in discriminability and bias between the first and second sets of ten problems completed by each participants. It seems clear that there is no evidence of change in either discriminability or bias. The log Bayes factors comparing the change model to one that assumes no change support this conclusions, showing evidence in favor of the no-change model of 1.93 for discriminability and 1.65 for bias.

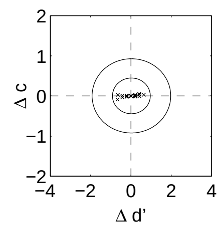

Figure 9. Change in individual discriminability and bias between the first 10 and second 10 problems completed. The markers show the expected change in discriminability ($\Delta d'$) and change in bias (Δc) for each participant between the first and second blocks of 10 problems. Superimposed are the 50% and 95% contours for the expected Gaussian distribution over the collection of differences.

Discussion

Individual Decision-Making

Our data for individual decision-making on the optimal stopping problem replicate all of the important findings that made group decision-making on the problem interesting. The raw data analysis in Figure 3 and the hierarchical signal detection analysis in Figure 5 both show large individual differences. The within-participants comparison of the first half or individual trials against the second half, as shown in Figure 9, shows no evidence of learning. And there is clear evidence that individual tend to make a choice too early in the sequence. Using signal detection theory, this can be seen most clearly in Figure 5, which shows that the bias for the individuals clearly errs on the lenient side of optimal decision-making.

Group and Individual Performance

Previous empirical findings for group decision-making on cognitive tasks have found considerable evidence that groups, typically with sizes between three and seven, rarely outperform their best members (see, for example, the reviews of Hastie & Kameda, 2005; Kerr & Tindale, 2004). Our data, in contrast, provide intriguing evidence that, under various circumstances, and to various extents, group decision-making can correct an individual bias of choosing too early in the optimal stopping problem.

This is clear in the analysis of the raw data in Figure 3, particularly when measuring participants' decision-making with respect to optimal processes rather than chance-influenced outcomes. In this way, we observe an improvement in discriminability for some group decision-making conditions over individual decision-making, to the extent that some groups clearly out-perform their best member. Hastie and Kameda (2005) suggest those examples showing superior group-decision typically use tasks in which different group members having different pieces of relevant information, or allowing one or more individuals in a group has the opportunity to convince the others of the 'correctness' of their decision. Those explanations are clearly not applicable here. An obvious difference between our study and many previous ones (as reviewed, for example, by Kerr & Tindale, 2004), is that group members all had exactly the same information available, and interacted only in the most limited of ways, by viewing each others accept or reject member decisions. These characteristics of the task preclude information pooling, and also do not support any deliberation process.

Insights from Signal Detection Analysis

Our use of hierarchical signal detection theory to analyze the entire sequence of accept and reject decisions provides a series of useful insights into the how the group conditions differ. In particular, it is able to isolate where in the decision-making process group decision-making diverges from being simply the combined decisions of a collection of independent individuals.

It is clear that in the consensus condition that member decisions are significantly more conservative than those made by the same participants as individuals. In addition, the consensus condition is inherently conservative, since it requires all members of the group to agree on an accept decision. Taken together, the left panel of Figure 5 shows that group decision-making is now essentially unbiased, and with improved discriminability.

In the leadership condition, member decisions are again significantly more conservative than the individual decisions. After these member decisions are viewed, the analysis of changes in Figure 8 shows the leader is sometimes further discouraged from their member accept decision, and this change is always a good one. Taken together with the initially more conservative member decision, the right panel of Figure 5 shows that these trends make the leadership group decisions much less biased, and also show improved discriminability.

In the majority condition, however, the behavior is quite different. There is no evidence that member decisions are different from individual ones, nor, indeed, that the group decisions differ in discriminability or bias from the individual decisions. In this sense, in the majority condition, the group behaves as a collection of individuals, whereas the consensus and leadership groups behave differently from the sum of their individual parts.

Our findings suggest, though, that where consensus and leadership group decision-making differs from individual decision-making is not where it might have been predicted. A straightforward prediction would be that individual and member decisions would be extremely similar, since the same information is available to the decision-maker in both circumstances, but that revised member decisions might be different, because of the additional information provided by seeing the recommendation of other group members. Our analysis makes very clear, however, that it is at the member stage that decisions differ, and relatively few revisions are made from that point onwards. This makes it difficult to explain the large changes in group decision-making in term of group polarization effects that have been a central focus in social group decision-making (e.g., Moscovici & Zavolline, 1969).

Accountability in Group Decision-Making

The difference between individual and member decisions in general, and the different (and inferior) behavior of the majority condition are interesting, and requires some explanation. This is particularly true since there is some evidence and advocacy[1] for the effectiveness of a majority rule in the existing literature (e.g. Hastie & Kameda, 2005; Sorkin, Hays, & West, 2001). The basic theoretical idea is that majority rules have the attraction of serving to amplify moderately correct individual decisions, especially in cases where the individual decisions are not strongly correlated.

One possible reason for this is that majority condition is the only one in which a member's decision is not necessarily directly responsible for a group decision. The leader's decision is the group decision, and it seems likely other members assume the leader will scrutinize their recommendation. In the consensus condition, all members must agree, and so everybody is directly accountable for an accept decision. In the majority condition, in contrast, the responsibility for both accept and reject decisions by the group can only be attributed to a collection of group members, and never to one individual.

Perhaps this lack of direct accountability is the reason the majority condition seems to differ from the other two. Such a line of argument seems related to the issue of group motivational gains, where group members exert greater effort than as individuals. Existing demonstrations of this effect (see Kerr & Tindale, 2004, p. 628, for an overview) typically involve different group decision-making situations, of a more inherently social nature. Nevertheless, at least one element believed to be important in these situations, that of social comparison, seems likely to be present in our experimental procedure. In group decision-making, member decisions are effectively individual decisions that will be seen by others. It is especially interesting, therefore, that when member decisions must coincide with group-decision in the consensus condition, or must be evaluated by a leader, people become

[1] Sorkin, Shenghua, and Itzkowitz (2004) advocated consensus group decision-making, but for the very different circumstance involving extensive information-sharing and deliberation.

more conservative, but the mere visibility of a member decision in the majority decision does not produce the same change.

Conclusion

We have presented an analysis of group decision-making, under three different decision-making conditions, on a well-controlled and easily measured optimal stopping task for which there are stable individual differences. Our primary finding is that, in the group setting, the decisions of individuals, for this task at least, are quite different from those they supplied as individuals, under conditions where their initial decision can be accountably linked to the decision of the group. This is, perhaps a surprising finding, especially given the fact that our participants had no interaction with one another in revising their decisions, and, in fact, were socially isolated from other group members, and that the task dealt with abstract stimuli in a mathematically described task. It may be the case, therefore, that the effect we observed is a pervasive one across more real-world stimuli and social settings. If true more generally, our findings suggest that theories and models of group decision-making on even abstract cognitive tasks need to focus not only on issues of information pooling and deliberation, but also on the latent effects merely being in a group have upon the decision-making of individuals.

Acknowledgments

We thank Bill Batchelder, Gary Ewing, and Michael Webb for comments, and James Campbell and Chrisi Lambos for assistance with data collection. MDL acknowledges the support of Australian Research Council grant DP0451793.

References

Bearden, J. N., Murphy, R. O., & Rapoport, A. (2005). A multi-attribute extension of the secretary problem: Theory and experiments. *Journal of Mathematical Psychology*, *49*, 410–422.

Campbell, J., & Lee, M. D. (in press). The effect of feedback and financial reward on human performance solving 'secretary' problems. In *Proceedings of the 28th annual conference of the cognitive science society.* Mahwah, NJ: Erlbaum.

Corbin, R. M., Olson, C. L., & Abbondanza, M. (1975). Contest effects in optimal stopping rules. *Organizational Behavior and Human Performance*, *14*, 207–216.

Cox, J. C., & Oaxaca, R. L. (1992). Direct tests of the reservation wage property. *Economic Journal*, *102*, 1423–1432.

Dudey, T., & Todd, P. M. (2001). Making good decisions with minimal information: Simultaneous and sequential choice. *Journal of Bioeconomics*, *3*(2–3), 195–215.

Ferguson, T. S. (1989). Who solved the secretary problem? *Statistical Science*, *4*(3), 282–296.

Gigone, D., & Hastie, R. (1997). Proper analysis of the accuracy of group judgments. *Psychological Bulletin*, *121*(1), 149–167.

Gilbert, J. P., & Mosteller, F. (1966). Recognizing the maximum of a sequence. *American Statistical Association Journal*, *61*, 35–73.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Hastie, R. (1986). Experimental evidence on group accuracy. In B. Grofman & G. Owen (Eds.), *Information pooling and group decision making* (pp. 129–157). Greenwich, CT: JAI Press.

Hastie, R., & Kameda, T. (2005). The robust beauty of majority rules in group decisions. *Psychological Review*, *112*(2), 494–508.

Kahan, J. P., Rapoport, A., & Jones, L. V. (1967). Decision making in a sequential search task. *Perception & Psychophysics*, *2*(8), 374–376.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773-795.

Kerr, N. L., & Tindale, R. S. (2004). Group performance and decision making. *Annual Review of Psychology*, *55*, 623–655.

Kogut, C. A. (1990). Consumer search behavior and sunk costs. *Journal of Economic Behavior and Organization*, *14*, 381–392.

Lee, M. D. (2006). A hierarchical Bayesian model of human decision-making on an optimal stopping problem. *Cognitive Science*, *30*, 555-580.

Lee, M. D., O'Connor, T. A., & Welsh, M. B. (2004). Human decision making on the full-information secretary probllem. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th annual conference of the cognitive science society* (pp. 819–824). Mahwah, NJ: Erlbaum.

Moscovici, S., & Zavolline, M. (1969). The group as a polarizer of attitudes. *Journal of Personality and Social Psychology*, *12*, 125–135.

Rapoport, A., & Tversky, A. (1970). Choice behavior in an optional stopping task. *Organizational Behavior and Human Performance*, *5*, 105–120.

Rouder, J. N., & Lu, J. (2005). An introduction to Bayesian hierarchical models with an application in the theory of signal detection. *Psychonomic Bulletin & Review*, *32*, 573–604.

Seale, D. A., & Rapoport, A. (1997). Sequential decision making with relative ranks: An experimental investigation of the "Secretary Problem". *Organizational Behavior and Human Decision Processes*, *69*(3), 221–236.

Seale, D. A., & Rapoport, A. (2000). Optimal stopping behavior with relative ranks. *Journal of Behavioral Decision Making*, *13*, 391–411.

Simon, H. A. (1976). From substantive to procedural rationality. In S. J. Latsis (Ed.), *Method and appraisal in economics* (pp. 129–148). London: Cambridge University Press.

Sorkin, R. D., Hays, C. J., & West, R. (2001). Signal-detection analysis of group decision-making. *Psychological Review*, *108*(1), 183–203.

Sorkin, R. D., Shenghua, L., & Itzkowitz, J. (2004). Group decision and deliberation: A distributed detection process. In D. Koehler & N. Harvey (Eds.), *Handbook of judgment and decision making.* New York: Oxford University Press.

Zwick, R., Rapoport, A., Lo, A. K. C., & Muthukrishnan, A. V. (2003). Consumer sequential search: Not enough or too much? *Marketing Science*, *22*(4), 503–519.