AFRL-RH-WP-TP-2008-0004

Signal Detection Analysis of Computer Enhanced Group Decision Making Strategies

Greg C. Elvers
University of Dayton
Dayton OH 45469

November 2007
Interim Report for August 2006 to August 2007

Air Force Research Laboratory
Human Effectiveness Directorate
Warfighter Interface Division
Collaborative Interfaces Branch
Wright-Patterson AFB OH 45433

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the 88th Air Base Wing Public Affairs Office and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

TECHNICAL REVIEW AND APPROVAL

AFRL-RH-WP-TP-2008-0004

THIS TECHNICAL REPORT HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION.

FOR THE DIRECTOR

//signed//   **//signed//**
Benjamin A. Knott   Daniel G. Goddard
Program Manager   Chief, Warfighter Interface Division
Collaborative Interfaces Branch   Human Effectiveness Directorate

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE		Form Approved OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Service, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)* November 2007	2. REPORT TYPE Interim	3. DATES COVERED *(From - To)* August 2006 – August 2007

4. TITLE AND SUBTITLE Signal Detection Analysis of Computer Enhanced Group Decision Making Strategies	5a. CONTRACT NUMBER FA8650-06-C-6738
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER 62202F
6. AUTHOR(S) Greg C. Elvers	5d. PROJECT NUMBER 2313
	5e. TASK NUMBER HC
	5f. WORK UNIT NUMBER 55

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Dayton Dayton OH 45469	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Materiel Command Air Force Research Laboratory Human Effectiveness Directorate Warfighter Interface Division Collaborative Interfaces Branch Wright-Patterson AFB OH 45433-7022	10. SPONSOR/MONITOR'S ACRONYM(S) AFRL/RHCP
	11. SPONSORING/MONITORING AGENCY REPORT NUMBER AFRL-RH-WP-TP-2008-0004

12. DISTRIBUTION AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

88 ABW PA Cleared 3/11/2008, WPAFB-08-0871.

14. ABSTRACT
Three experiments were performed to test Sorkin's (2001) signal detection model of group decision making. The first two experiments compared various sequencing protocols which determine the order in which group members speak during group deliberation. The third experiment also compared these sequencing protocols but used correlated information sources so that some group members shared some information sources. It was predicted that the System for Optimally Rapid Collaboration (SORC) sequencing rule would produce the most rapid and accurate group decisions. While the results are mixed, there is some evidence that the SORC method is more accurate and in some situations more rapid than a random sequencing protocol.

15. SUBJECT TERMS
System for Optimally Rapid Collaboration (SORC), Signal Detection, Group Decision Making

16. SECURITY CLASSIFICATION OF: Unclassified			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Benjamin Knott
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U	SAR	38	19b. TELEPONE NUMBER *(Include area code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI-Std Z39-18

This page intentionally left blank.

Sorkin and his colleagues (Sorkin, Kantowitz, & Kantowitz, 1988; Elvers & Sorkin, 1989; Sorkin, Mabry, Weldon, & Elvers, 1991) have a long history of using signal detection theory to model decision making in individuals. Recently Sorkin and his colleagues (Sorkin, 2001; Sorkin, Luan, & Itzkowitz, 2004; Sorkin, 1998; Sorkin & Dai, 1994; Sorkin, Hays, & West, 2001; Sorkin, West, & Robinson, 1998) have developed and tested a normative model of group decision making based on signal detection theory. This model is summarized in Figure 1.

Because the model is based on signal detection theory, two of its key parameters are a measure of sensitivity, d', and a measure of response bias, β. A sensitive group member, or a sensitive group is one that usually responds that a signal event has occurred when in fact it has occurred, and rarely responds that a signal event has occurred when in fact no signal event has occurred. Sensitive individuals or groups will have a large value of d'. Typical values of d' range from 0, which indicates chance level performance, to approximately 4, which is a very sensitive group member or group. β is a measure of how willing a group member or a group is to state that a signal event has occurred. When the logarithm of β is less than 0, the group member or group is said to be liberal, and requires little evidence that the signal event occurred before they are willing to state that a signal event did occur. When the logarithm of β is larger than 0, the group member or group is said to be conservative, and requires much evidence that the signal event occurred before they are willing to state that the signal event did occur. The values of d' and β can be calculated given the probability of a hit (the group member or group states that the signal event occurred when it did) and the probability of a false alarm (the group member or group states that signal event occurred when it did not.)

The model assumes that each member of the group receives information about whether a signal event has occurred or not. This information is noisy such that it does not perfectly predict whether the signal event has occurred. The noise comes from two sources: a common noise source for all the members of the group, and an independent noise source that is unique for each member of the group. By changing the ratio of the independent and common noise sources, the model can accommodate situations in which each group member has unique information (only the independent noise sources are used) to situations in which each group member has the same information (only the common noise source is used) to anything in-between. The model can also handle situations in which some group members receive more information than others or some group members receive higher quality information than others.

Based on his or her information, each group member is assumed to calculate the likelihood ratio that the information came from the signal or no-signal event. This likelihood ratio is compared to the person's response criterion, β, for deciding that a signal event has occurred. The response criterion can be conservative, in which case the group member requires much evidence of the signal event before he or she is willing to state that the signal event has occurred, or liberal, in which case the group member requires little evidence of the signal event before he or she is willing to state that the signal event has occurred, or anywhere in between. The group member decides to vote that the signal event has occurred if and only if the natural logarithm of the likelihood ratio is greater than or equal to the natural logarithm of β, the response criterion.

Next, the model assumes that a secret ballot by the group members occurs. If the number of yes, a signal event has occurred, votes exceeds a criterion (e.g. a simple majority, a super majority, unanimity) then, according to the model, a group decision has been made that the

signal event occurred. Likewise, if the number of no, a signal event has not occurred, votes exceeds a criterion, then a group decision has been made that the signal event did not occur.   If the group did not reach consensus, deliberation occurs.

During deliberation, the vote of each group member is sequentially presented to the group. The group member's performance history, sensitivity, and response criterion may also be shared with the group. Based on this information, the group members revise their likelihood ratios and βs in a Bayesian manner. Another secret ballot is cast, and if group consensus has not been reached, deliberation continues until either consensus is reached or the votes of each group member have been shared with the group.

The order in which each group member's vote and related information (history, sensitivity, etc.) is sequentially presented to the group during deliberation is determined by the response sequence protocol. Different response sequence protocols could lead to differences in group decisions and length of deliberation.   For example, the sequencing of vote and information could be based on speed of response to the initial presentation of the information. Those group members who respond quickly may be more confident of their decisions than group members who respond more slowly. If confidence is partially determined by the quality of the information that a group member receives, then presenting the votes of group members who respond quickly early in the deliberation process should lead to more accurate group decisions and shorter deliberations than if the same, high quality, information was presented later in the deliberation. As another example, the response sequence protocol could be based on how unique a group member's information is. Group members who have information that others do not may provide more unique information to the deliberation process than those whose information overlaps with others. Thus, presenting the votes of group members who have unique information

3

early in the deliberation process should lead to more accurate group decisions and shorter deliberations than if the same, highly unique, information was presented later in the deliberation.

The model is normative in that it can be used to make predictions about how an ideal group behaves. Signal detection theory also provides a measure of efficiency, η, which provides a measure of the extent to which a real group performs below the ideal (Tanner & Birdsall, 1958):

$$\eta = \left(\frac{d'_{observed}}{d'_{ideal}} \right)^2$$

Because the model is normative, one can predict how an ideal group would make its decisions based on, for example, how inter-correlated the group members' information is, the quality of the group members' information, the proportion of votes required to reach consensus, the sensitivity of each group member, group size, and the response sequence protocol.

Empirical tests of the model are typically performed by presenting each group member with a vector of numbers which have been sampled from one of two Gaussian distributions – the signal event distribution or the no signal event distribution. These numbers form the basis from which each group member decides whether or not the signal event has occurred or not. In general, the larger the values are, the more likely that a signal event has occurred. The smaller the values are, the less likely that a signal event has occurred. By varying the difference between the means of the two distributions, and the standard deviation of the distributions, one can vary the quality of the information that a given group member receives. Quality is directly proportional to the difference of the means of the distributions, and inversely proportional to the standard deviations of the distributions. Quality of information can also be manipulated by varying the length of the vector – the larger the number of samples drawn from a distribution, the greater the information about the signal event will be.

Itzkowitz (2005) compared ideal to observed group decision making performance using five different response sequence protocols. The five response sequence protocols were based on group member's response times (from fastest to slowest), group member's ability (from largest to smallest d'), group member's ability designed to prolong deliberation (from smallest to largest d'), a measure of group member confidence (from largest to smallest differences between the likelihood ratio and β), and a random sequence as a control. For all response sequence protocols, ideal and observed decision accuracy closely match, but observed accuracy was slightly less than the ideal. The response sequence protocol designed to prolong deliberation (from smallest to largest d') had the highest level of group accuracy while the remaining response sequence protocols (response time, ability, confidence, and random) were approximately equally accurate.

The length of deliberation also yielded similar results between the ideal and observed group decisions. The observed groups reached consensus, on average, more quickly than the ideal groups did. The response sequence protocol designed to prolong deliberation was successful, in that it had a larger number of votes during deliberation than any of the other response sequence protocols, although none of the response sequence protocols were reliably different from each other.

Sorkin, West, and Robinson (1998) tested the model's ability to predict group decision making without deliberation for different consensus criteria. The consensus criterion is the proportion of the group who must vote that the signal event occurred in order for the group to vote that the signal event occurred. In one condition a simple majority of yes votes was needed for the group to decide that the signal event had occurred. In another condition, three-quarters of the group members had to vote yes. In the final condition, the vote had to be unanimous for the group to decide that the signal event had occurred. Sorkin's model predicted that as the

proportion of votes required for the group to reach consensus increased from 0.5 to 1.0 that the group should become less and less accurate in their decision. This arises because an ever decreasing number of group members can veto the rest of the group's decisions. The model also predicted that as the proportion of votes required for the group to reach consensus increased from 0.5 to 1.0 that the group should become increasingly conservative in their decision. This follows from the preceding prediction – if a smaller number of group members can veto the overall group, then an increasing number of group members must have good evidence of the signal event before the group will vote that the signal event occurred. Finally, Sorkin, West, and Robinson predicted that some group members would become increasingly liberal in their individual votes to counteract the increasing conservatism that should arise as the number of votes required for consensus increases. If some of the group members realize that the group is becoming more conservative as the consensus criterion became more stringent, they could partially counteract the conservatism of the group by becoming more liberal with their individual decisions. Each of these predictions was supported by empirical tests of the model.

Sorkin, Hays, and West (2001) also used the model to make predictions about group decision making. Group performance should increase as the size of the group increases, but the extent of the increase should depend on the extent of inter-correlation among the information that the individual group members receives. When the information is independent, performance should increase with the square root of group size. When the information is even moderately correlated ($\rho > .25$) the model predicts that much of the advantage of group size will disappear. This follows from the model because each additional member of the group receives additional, perhaps unique information which could aid in the group decision making process. However, the quantity of additional, unique information is reduced per additional group member when the

6

correlation between the group member's information increases. In general, the empirical data agreed with the predictions.

Sorkin, Hays, and West (2001) also investigated how the information was integrated. That is, how strongly did another group member's information influence the revision to a given group member's vote? Sorkin, Hays, and West argue that the weights given to another's vote should be proportional to the difference between the mean of the distribution from which signal event information is drawn for the individual and the distribution from which no signal event information is drawn for the individual. The weights should be inversely proportional to the variance of the distributions. This follows from the fact that the quality of the information increases as the difference between the means of the distributions increases, and decreases as the variance of the distributions increases.

Using the conditional on a single stimulus (COSS) method of Berg (1989, 1990), Sorkin, Hays, and West (2001) determined the weight that each group member assigned to the other group member's votes. Berg's COSS method basically involves correlating the group's decision with the mean value of the information that an individual group member received for each member of the group. If a group member's information is highly correlated with the group's decision, then that group member's vote is very influential when the group makes its decision and the weight given to that group member is relatively high. If a group member's information is not correlated with the group's decision, then that group member's vote does not influence the group's decision and the weight given to that group member is close to 0. Finally, the weights are transformed to efficiencies using Berg's (1990) weighting efficiency formula:

$$\eta_{\text{weight}} = \left(\frac{d'_{\text{actual-weights}}}{d'_{\text{optimal-weights}}} \right)^2$$

7

Where d'$_{actual-weights}$ is value of d' that would occur if the only inefficiency was due to using the observed weights instead of the optimal weights, and d'$_{optimal-weights}$ is the value of d' that would have occurred if the ideal weights had been used. In general, the weighting efficiencies were large indicating that the observers behaved in a manner consistent with the predictions of the model. However, there were some exceptions to the general rule. The group member who cast the group's decision tended to overweigh, compared to optimal weights, his or her own information. Also, as group size increased, weighting efficiencies tended to decrease more rapidly than predicted by the model. Sorkin, Hays, and West attribute this, at least partially, to the fact that the larger groups tended to rush through deliberation. Rushing through deliberation, especially for larger groups that would require additional time for each additional group member to announce their vote, would have maximized the hourly wage that the group members earned, as they were paid a base rate, plus a bonus based on the number and quality of group decisions made. Sorkin, Hays, and West also speculate that social loafing may have played a role – with larger groups, each individual member may have worked less hard.

Based on the above research, Sorkin's model has demonstrated its ability to accurately model the human decision making process. One of the least investigated, but most promising components of the model, the response sequencing protocol, offers the possibility of improving the accuracy of the group decision while simultaneously decreasing the length of deliberation. A simultaneous increase in accuracy and decrease in length of deliberation would be highly desirable in many military and non-military situations where decision time is extremely limited and errors can have a huge negative impact.

Sorkin (personal communication, 22 March 2006) has devised a new, proprietary response sequence protocol called the System for Optimally Rapid Collaboration (SORC)

protocol. Given the individual group member's initial votes and information, the SORC method determines every possible sequence of presenting the individual group member's votes to the group, and using Bayes' Theorem to update each individual's vote, selects the sequence that yields the shortest series of votes which leads to the optimal decision.

Sorkin (personal communication, 22 March 2006) has performed some preliminary Monte Carlo simulations of the SORC protocol based on his model which indicate that the SORC protocol may enhance group decision making performance in terms of both accuracy of the group decision, and a decrease in the time spent deliberating. While the simulations are encouraging, the SORC protocol needs to be empirically tested. The primary goal of the this research is to evaluate the effectiveness of the SORC protocol compared to other response sequence protocols that have yielded good (but not ideal / optimal) performance (both accuracy and length of deliberation) in previous studies.

Experiment 1

The first experiment was meant to be a simple test of the SORC protocol, software, instructions and procedures. On each trial the computer would randomly, with equal probability, decide whether a signal or noise event had occurred. The computer then generated nine values for each group member. If a signal event occurred, the values tended to be higher than if a noise event occurred. Each group member looked at their individual information displays that displayed these values on nine gauges (see Figure 2). Based on the displayed values, each group member decided whether a signal or noise event was more likely. If these initial votes were not unanimous, the group deliberation began. During group deliberation, one of the group member's votes and other information about that group member (accuracy and bias) was shown to the rest of the group. The rest of the group updated their own vote based on the initial information and

the new information that was just shared. The deliberation process continued until the group

either voted unanimously or the individual votes of all group members had been shared.

Method

Participants

Two groups of six college students participated in the study. The first group consisted of

five females and one male. The second group consisted of six females. On any given day, if all

six group members showed up, one was randomly dismissed and the rest formed that day's

group. They were paid based on their individual and group performance at the decision making

task. The mean hourly rate was approximately $9.

Design

The group deliberation sequencing protocol was manipulated within-subjects with one of

four values: random, maximum confidence, maximum d', and SORC. With the random

sequencing protocol, the group member whose information was shared next during group

deliberation was selected at random. With the maximum confidence sequencing protocol, the

distance of each group member's information from the signal or noise mean (depending on

whether a signal or noise event was occurring on a particular trial) was calculated and sorted

from highest to lowest. This determined the order in which each individual's responses were

shared with the group during group deliberation. In the maximum d' sequencing protocol, the

optimal d' of each group member was calculated (based on standard deviation used to determine

the values of their gages), sorted from highest to lowest, and the group member with the highest

optimal d' whose information had not yet been shared was shared next. With the SORC

sequencing protocol, all possible (5! = 120 on the first group vote) combinations of sharing

orders were considered. For each possible order, each group member's response was estimated

using Sorkin's model of group decision making and Bayes' theorem to update each group member's response. The sequence which led to the shortest, correct answer was selected. If more than one sequence led to the shortest, correct answer, one of the shortest, correct sequences was selected at random. If none of the sequences led to the correct answer, one of the shortest sequences that led to the incorrect answer was selected at random.

Each group member's initial vote was recorded. If the group did not initially vote unanimously, the individual votes during group deliberation was also recorded. The number of group deliberation votes needed to reach a unanimous decision and the accuracy of that decision was recorded.

Apparatus

Six PC compatible computers were connected in a local area network. One computer acted as the server computer which directed the other five computers to display the stimuli. The other five computers were the client computers which displayed stimuli and collected each group member's responses.

Procedure

The experiment consisted of three parts: individual practice, group practice, and group data collection. During individual practice each block of trials started with an information screen that informed each group member how much he or she would earn for making each type of response; that is, they were shown the signal detection theory payoff matrix. The payoff matrices resulted in the optimal group members having a response bias (log β) of -0.22 (liberal), 0.00 (unbiased) or 0.22 (conservative.) On each trial each group member saw a set of nine analog gauges like those in Figure 2. The server determined the values of the nine gauges in the following manner. First, the server decided whether a signal plus noise event or a noise only

11

event occurred with equal probability. If the signal plus noise event occurred, the server generated nine normally distributed random numbers with a mean of 5 for each of the five group members. If the noise only event occurred, the server computer generated nine normally distributed random numbers with a mean of 4 for each of the five group members. The standard deviation of the normal distribution was gradually increased from 1 to 3 during the first two days of practice. These random values then were displayed on the analog gauges to the group members. The display duration was decreased from 3000 ms to 300 ms during the first two days of practice. After displaying the gauges, they were replaced with the individual response screen as shown in Figure 3. Each group member, without communicating to the other group members, made two responses. First, each clicked on the confidence rating bar to indicate how confident he or she was in his or her decision. Second, each group member clicked on either the signal button or the noise button to indicate whether they believed that a signal plus noise event had occurred (relatively higher values on the gauges) or a noise only event occurred (relatively lower values on the gauges.) After all group members had made their responses, each received individual feedback about the accuracy of her response. Each group member received over 900 trials of individual practice with display characteristics (stimulus duration and standard deviations of the gauge values) identical to those that would be used during data collection.

The second part of the study involved group practice. Each trial of group practice started in the same way as the individual practice – the server determined which type of event occurred, generated appropriate values for the gauges which were then shown to each group member who made a confidence rating and an individual decision about whether signal plus noise or noise only occurred. If the individual votes were not unanimous, group deliberation began. During group deliberation each group member saw a display similar to Figure 4. The group deliberation

screen showed several pieces of important information: the group member's confidence in their

individual vote, the performance of each group member if they were performing optimally (as

specified by signal detection theory), and the previous response of one of the group members.

As shown in the middle of Figure 4, the ideal performance of each group member was shown by

four bars. Two were turquoise and indicated the optimal hit (leftmost bar) and correct rejection

(rightmost bar) rates of the group member. Two were brown and indicated the optimal miss (the

second bar from the left) and false alarm rates (the second bar from the right) rates of the group

members. The length of each bar, in arbitrary units, was also displayed within each bar.

The longer each bar was, the more likely the group member, if optimal, would make that

type of response. The longer the sum of the turquoise bars (the first, hit and fourth, correct

rejection bars), the more accurate the optimal group member tended to be. The length of the bars

also indicated whether the group members were biased in their responses or not, based on the

payoff matrix that each received. If the sum of the lengths of the hit (first) and false alarm (third)

bars were longer than the sum of the length of the correct rejection (fourth) and miss (second)

bars, then the optimal group member's response was liberally biased (they would tend to make

more signal responses.) If the pairs were equal in length, the optimal group member's response

was not biased. If the sum of the lengths of the hit and false alarm bars were shorter than the

sum of the lengths of the correct rejection and miss bars, then the optimal group member's

response was conservatively biased (then would tend to make fewer signal responses.)

The group members were told the above information and told to utilize the information

in the display to update their decision as to whether a signal plus noise or noise only event

occurred. The group members were told that there were no protocols for updating their

decisions, but that several factors should enter their decision to either stay with their original

13

response or to switch responses: the group member's original vote and confidence, the group

member optimal performance, the previous response of one of the group members, and the

optimal performance of that group member. Instructions about how each of these items should

influence the group member's vote were given. Next, each group member, without

communicating with the other group members, clicked on either the signal or noise button to

indicate their current decision. If the decisions were not unanimous, the process repeated either

until a unanimous decision was reached or each group member's information had been shared.

At the end of the deliberation, or immediately after the individual decisions were made if

they resulted in a unanimous decision, each group member received feedback about the accuracy

of his or her initial vote and the group vote, how much money she or he earned on that trial, and

how much money she or he had earned in the current session. Each group member received over

900 trials of group decision making practice using the random sequencing protocol.

The third part of the study consists of the data collection. From the group members'

standpoint, it was identical to the second part of the study, group practice. The sequencing

protocol was manipulated across sessions using a reverse counterbalancing scheme. Group one

experienced 960 trials for each of the four sequencing protocols. Group two experienced 990

trials for each of the four sequencing protocols.

Results

Table 1 shows the mean number of group votes that occurred before the final decision

was reached (either through consensus or after all group members had spoken and no consensus

was reached, in which case the majority vote was considered the group's final vote.) For group

1, none of the sequencing protocols produced a reliable difference in the mean number of group

votes that occurred before the final decision was reached. For group 2, the maximum confidence

sequencing protocol produced the largest number of votes before reaching the final decision. The maximum d' and SORC sequencing protocols resulted in a similar number of group votes prior to the final decision, while the random sequencing protocol resulted in smallest number of group votes prior to the final decision. Table 1 also shows the standard deviation of the group votes, and the standard error of the mean of the group votes. Finally, Table 1 shows the probability that each pair of means was sampled from the same parent distribution.

Table 2 shows the group decision d' and ß for each sequencing protocol. These results show that for both groups d' and ß were relatively unaffected by the sequencing protocol.

For group 1 these results indicate that the sequencing protocol did not influence the number of votes to the final decision, d' or ß. For group 2, there is a slight hint of a speed-accuracy trade off for some of the sequencing protocols – the random sequencing protocol lead to the fastest decisions (based on the number of group votes to the final decision), but also was the least sensitive. The maximum d' sequencing protocol lead to an intermediate speed of decision and an intermediate sensitivity. The maximum confidence sequencing protocol also had intermediate sensitivity, but was the slowest of the sequencing protocols for reaching the final decision. Finally, the SORC sequencing protocol had intermediate speed of decision and was the most sensitive.

Experiment 2

Experiment 2 was a replication of the first experiment with minor changes to the placement of the bars in the group deliberation screen to better emphasis the response bias of the optimal group members and more detailed instructions.

Method

Participants

15

Two groups of college students participated in the study. The first group consisted of

five females. The second group consisted of four females and two males. On any given day, if

all six group members showed up for group two, one was randomly dismissed and the rest

formed that day's group. We could not recruit a sixth group member for the first group, so each

group member always participated. They were paid based on their individual and group

performance at the decision making task. The mean hourly rate was approximately $9.

Design

The design was similar to experiment one with the following exceptions. The order of

the bars on the group deliberation screen where changed to make the response bias of the optimal

group member more perceptually salient (see Figure 5.) The first and fourth bars continued to

represent the hit and correct rejection rates of the optimal group member. The second and third

bars' positions were switched so that the second bar represented the false alarm rate and the third

bar represented the miss rate. If the length of the first and second bars were longer than the

length of the third and fourth bars, the optimal group member was liberally biased, while if the

first and second bars where shorter than the third and fourth, the optimal group members was

conservatively biased. Second, only three response sequencing protocols were tested: random,

maximum d' and SORC.   Third, the instructions were modified to more fully describe the types

of information that the group members should use during group deliberation. Finally, each group

member experienced 630 trials for each of the three sequencing protocols.

<center>*Results*</center>

Table 3 shows the mean number of group votes that occurred before the final decision

was reached (either through consensus or after all group members had spoken and no consensus

was reached, in which case the majority vote was considered the group's final vote.) For group

<center>16</center>

3, the random and maximum d' sequencing protocols resulted in a similar number of group votes prior to the final decision, while the SORC protocol resulted in a smaller number of group votes prior to the final decision. For group 4, the random and SORC sequencing protocols resulted in a similar number of group votes prior to the final decision, while the maximum d' sequencing protocol resulted in a smaller number of group votes prior to the final decision. Table 3 also shows the standard deviation of the group votes, and the standard error of the mean of the group votes. Finally, Table 3 shows the probability that each pair of means was sampled from the same parent distribution. For group 3 these results indicate that the SORC sequencing protocol lead to faster group decisions than the random sequencing protocol, and perhaps faster group decisions than the maximum d' sequencing protocol. For group 4 these results do not indicate a superiority of the SORC protocol for reaching rapid decisions.

Table 4 shows the group decision d' and ß for each sequencing protocol. These results show that for both groups the maximum d' and SORC sequencing protocols lead to better group decisions than the random sequencing protocol. Along with the results from Table 3, the SORC sequencing protocol seems to lead to the fastest group decisions and among the most accurate group decisions for group 3, and accurate decisions for group 4.

<div align="center">Experiment 3</div>

In both experiments one and two, the information that each group member received was independent of the other group member's information, except that the information that all group members saw was all drawn from the signal distribution or all drawn from the noise distribution. That is, it was as if each individual observed the same situation but did so in completely independently and derived unique information from the situation. Experiment three considers the case where some of the information is shared across group members.

<div align="center">17</div>

Method

Participants

One group of college students participated in the study (a second group will participate during the winter 2008 semester.) The group consisted of five females. They were paid based on their individual and group performance at the decision making task. The mean hourly rate was approximately $9.

Design

The design was similar to experiment two with the following exceptions. The gauges were correlated between two pairs of group members. For two of the group members, the values displayed on the leftmost four gauges were always identical. For another pair of group members, the values displayed on the fifth through eighth gauges were always identical. This information was conveyed to the group members on the group deliberation screen (see Figure 6) which included red arcs and numbers between each pair of group members. The red number near a given arc indicated how many gauges were identical for that pair of group members.

Results

Table 5 shows the mean number of group votes that occurred before the final decision was reached (either through consensus or after all group members had spoken and no consensus was reached, in which case the majority vote was considered the group's final vote.) For group 5, the random and SORC sequencing protocols resulted in a similar number of group votes prior to the final decision, while the maximum d' protocol resulted in a smaller number of group votes prior to the final decision. Table 5 also shows the standard deviation of the group votes, and the standard error of the mean of the group votes. Finally, Table 5 shows the probability that each pair of means was sampled from the same parent distribution. For group 5 these results indicate

that the maximum d' sequencing protocol lead to faster group decisions than the random sequencing protocol and the SORC sequencing protocol.

Table 6 shows the group decision d' and ß for each sequencing protocol. These results show that the maximum d' and SORC sequencing protocols lead to better group decisions than the random sequencing protocol. Along with the results from Table 5, the maximum d' sequencing protocol seems to lead to the fastest group decisions and the most accurate group decisions for group 5. The SORC sequencing protocol is as accurate as the maximum d' protocol is, but slightly slower.

<div align="center">Discussion</div>

The results do not strongly support the SORC sequencing protocol as a more rapid and accurate sequencing protocol compared to the maximum d' sequencing rule. In experiment one, the sequencing protocol did not affect the speed of decisions for group 1 and the SORC protocol had one of the lowest d's. For the second group of experiment one, the SORC method lead an intermediate speed of decisions and had the highest d'. In experiment two, the SORC protocol lead to the quickest and accurate decisions for group 3, but was slower and still accurate for group 4. In experiment three, the SORC protocol again lead to slower decisions that were accurate. Overall, the performance, in terms of both speed and accuracy, of the SORC protocol were approximately equal to those of the maximum d' rule which is much simpler computationally.

Several reasons might exist why the SORC sequencing protocol is not rapid and accurate in practice as it is in theory. The participants may not be able to effectively utilize all of the information that is present on the group deliberation screen in updating their decision. In the third experiment with five group members there are up to 47 pieces of information on the

display (prior estimate, last secret vote, four response parameters [hit, miss, correct rejection, false alarm rates] for each group member, accuracy information [length of turquoise bars] for each group member, response bias information [length of left two bars relative to the right two bars] for each group member, up to five prior shared opinions, 10 arcs indicating gauge intercorrelations, and the current shared vote.) The amount of information might be so great that the participants select only a small subject on which to base their revised votes. If they select information that is different from what the SORC sequencing protocol uses, then their decisions should be different from those predicted by the SORC protocol.

Another potential reason is that human participants may not act as Bayesian observers as assumed by Sorkin's (2001) model. While decision maker's behavior may approximate that predicted by Bayes' theorem, it is unlikely that the group members actually engage in the type of probabilistic updating of their likelihood ratios as mandated by Bayes' theorem.

As discussed by Sorkin, Hays and West (2001) also discussed the possibility of social factors, such as social loafing, in group decision making. Sorkin's model does not consider these factors. To the extent that social factors play a large role in group decision making, Sorkin's model will not accurately predict group decision making.

References

American Psychological Association (2002). Ethical principles of psychologists and code of

conduct. *American Psychologist, 57*, 1060-1073.

Itzkowitz, J. (2005). Improving performance of deliberative groups through changes in

organizational structure: A signal detection approach (Doctoral dissertation, University of

Florida, 2005). *Dissertation Abstracts International, 66*, 5116.

MacMillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.).

Mahwah, NJ: Lawrence Erlbaum.

Sorkin, R. D. (1998). Group performance depends on the majority rule. *Psychological Science,

9*, 456-463.

Sorkin, R. D. (2001). Signal-detection analysis of group decision making. *Psychological

Review, 108*, 183-203.

Sorkin, R. D., & Dai, H. (1994). Signal detection analysis of the ideal group. *Organizational

Behavior and Human Decision Processes, 60*, 1-13.

Sorkin, R. D., Hays, C. J., & West, R. (2001). Signal detection analysis of group decision

making. *Psychological Review, 108*, 183-203.

Sorkin, R. D., Kantowitz, B. H., & Kantowitz, S. C. (1988). *Human Factors, 30*, 445-459.

Sorkin, R. D., Luan, S., & Itzkowitz, J. (2004). Group decision and deliberation: A distributed

detection process. In D. J. Koehler, & N. Harvey (Eds.), *Blackwell handbook of judgment

and decision making* (pp. 464 – 484). Malden, MA: Blackwell Publishing.

Sorkin, R. D., Mabry, T. R., Weldon, M. S., & Elvers, G. C. (1991). Integration of information

from multiple element displays. *Organizational Behavior and Human Decision

Processes, 49*, 167-187.

Table 1

Number of Group Votes Needed To Reach the Final Decision for Experiment 1

Sequencing Protocol	Mean Number of Group Votes to Final Decision	Standard Deviation of Group Votes to Final Decision	$s_{\overline{X}}$ of Group Votes to Final Decision	p(two samples from same populations)		
				SORC	Max. Conf.	Max. d'
Group 1:						
Random	4.40	2.27	0.073	0.320	0.398	0.500
Maximum d'	4.40	2.26	0.073	0.320	0.398	
Maximum Confidence	4.37	2.31	0.075	0.418		
SORC	4.36	2.32	0.075			
Group 2:						
Random	3.69	2.23	0.092	0.016	0.001	0.018
Maximum d'	3.95	2.21	0.091	0.479	0.001	
Maximum Confidence	4.37	2.13	0.087	0.001		
SORC	3.96	2.24	0.092			

Table 2

d' and ß for each Sequencing Protocol for Experiment 1

Sequencing Protocol	Group 1		Group 2	
	d'	ß	d'	ß
Random	1.67	0.65	1.52	0.77
Maximum d'	1.82	0.71	1.61	0.61
Maximum Confidence	1.78	0.74	1.59	0.76
SORC	1.69	0.53	1.74	0.57

Table 3

Number of Group Votes Needed To Reach the Final Decision for Experiment 2

| | | | | p(two samples from same populations) | |
Sequencing Protocol	Mean Number of Group Votes to Final Decision	Standard Deviation of Group Votes to Final Decision	$s_{\overline{X}}$ of Group Votes to Final Decision	SORC	Max. d'
Group 3:					
Random	2.52	2.19	0.087	0.002	0.282
Maximum d'	2.45	2.21	0.088	0.002	
SORC	2.17	2.14	0.083		
Group 4:					
Random	3.33	2.13	0.085	0.433	< 0.001
Maximum d'	2.92	2.09	0.083	< 0.001	
SORC	3.35	2.18	0.087		

Table 4

d' and ß for each Sequencing Protocol for Experiment 2

	Group 3		Group 4	
Sequencing Protocol	d'	ß	d'	ß
Random	1.55	0.70	1.84	0.77
Maximum d'	1.95	0.48	2.09	0.69
SORC	1.82	0.58	2.08	0.70

Table 5

Number of Group Votes Needed To Reach the Final Decision for Experiment 3

Sequencing Protocol	Mean Number of Group Votes to Final Decision	Standard Deviation of Group Votes to Final Decision	$s_{\overline{X}}$ of Group Votes to Final Decision	SORC	Max. d'	p(two samples from same populations)
Group 5:						
Random	3.33	2.13	0.085	0.433	0.001	
Maximum d'	2.92	2.09	0.083	0.001		
SORC	3.35	2.18	0.087			
Group 6:						
Random						
Maximum d'						
SORC						

Note. Experiment 3 is still in progress. Data for group 5 has been collected, but data for group 6 will be collected during the winter 2008 semester.

Table 6

d' and ß for each Sequencing Protocol for Experiment 3

Sequencing Protocol	Group 5		Group 6	
	d'	ß	d'	ß
Random	1.84	0.77		
Maximum d'	2.09	0.69		
SORC	2.08	0.70		

Note. Experiment 3 is still in progress. Data for group 5 has been collected, but data for group 6 will be collected during the winter 2008 semester.

Figure Captions

Figure 1. Sorkin's model of group decision making is based on signal detection theory (from Sorkin & Dai, 1994.)
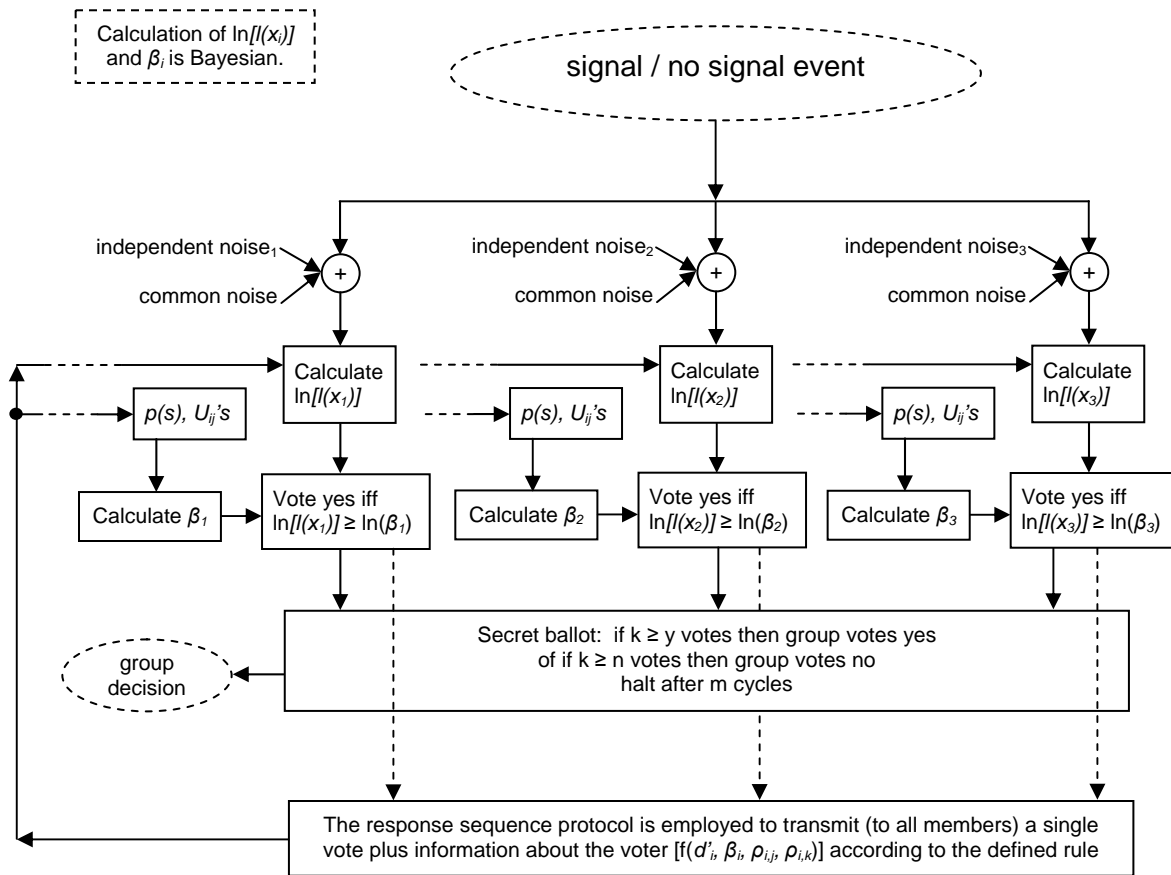
Figure 2. The information screen which is shown at the start of each trial. The gauge values tend to be higher when a signal event has occurred and lower when a noise event has occurred.
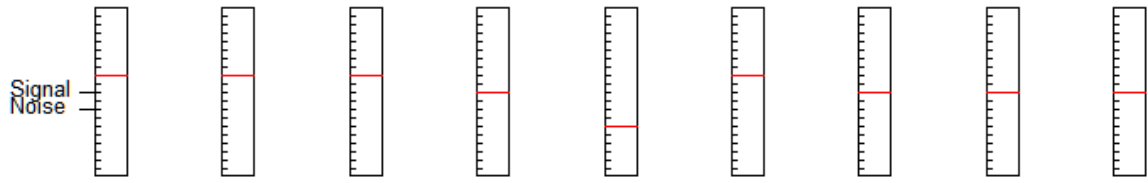
Figure 3. After seeing the information screen, each group member saw the individual response screen where each group member indicated how confident they were in their decision and whether they believed a signal or noise event was more likely.

Figure 4. Experiment 1's group deliberation screen shows the accuracy (length of the turquoise bars) and response bias (whether the sum of the first and third bars are longer or shorter than the sum of second and fourth bars) for each of the group members.

Figure 5. The modified group deliberation screen from experiment 2 more saliently shows the response bias of the optimal group members. The longer the left two bars are compared to the right two bars for each group member, the more liberally biased the optimal group member is.

Figure 6. Experiment 3's group deliberation screen is similar to experiment 2's screen except that red arcs indicating the amount of overlap or sharing between each group member's observations. In this example, group members 1 and 2 share four of their nine gauges – that is, four of their gauges showed identical values. Group members 1 and 4 did not have any gauge values in common as indicated by the 0 next to the red arc between those two group members.

Calculation of ln$[l(x_i)]$
and β_i is Bayesian.

signal / no signal event

independent noise$_1$ +   independent noise$_2$ +   independent noise$_3$ +

common noise   common noise   common noise

Calculate ln$[l(x_1)]$ Calculate ln$[l(x_2)]$ Calculate ln$[l(x_3)]$

$p(s)$, U_{ij}'s $p(s)$, U_{ij}'s $p(s)$, U_{ij}'s

Calculate β_1 Vote yes iff ln$[l(x_1)] \geq$ ln(β_1) Calculate β_2 Vote yes iff ln$[l(x_2)] \geq$ ln(β_2) Calculate β_3 Vote yes iff ln$[l(x_3)] \geq$ ln(β_3)

Secret ballot: if k ≥ y votes then group votes yes
of if k ≥ n votes then group votes no
halt after m cycles

group decision

The response sequence protocol is employed to transmit (to all members) a single vote plus information about the voter [f(d'_i, β_i, $\rho_{i,j}$, $\rho_{i,k}$)] according to the defined rule

30

Signal —
Noise —

Confidence Rating

Not Very Confident   Very Confident

Respond

| Noise | | Signal |

Your estimate was [slider]

Your last (secret) vote was: Noise

Member	Prior Shared Opinions	Average Performance

1 | 69 | 31 | 31 | 69 |

2 | 84 | 16 | 50 | 50 |

3 | 50 | 50 | 16 | 84 |

4 | 90 | 10 | 23 | 77 |

5 | 77 | 23 | 10 | 90 |

Member 4 is now sharing a new opinion: Signal

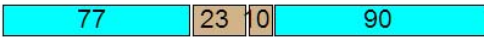The group has not yet reached consensus. Please vote again:

| Noise | Signal |

Your estimate was

Your last (secret) vote was: Noise

Member   Prior Shared Opinions   Average Performance

20

| 84 | 50 | 16 | 50 |

21

| 77 | 23 | 23 | 77 |

22

| 77 | 10 | 23 | 90 |

23

| 90 | 23 | 10 | 77 |

24

| 50 | 16 | 50 | 84 |

Member 21 is now sharing a new opinion: Signal
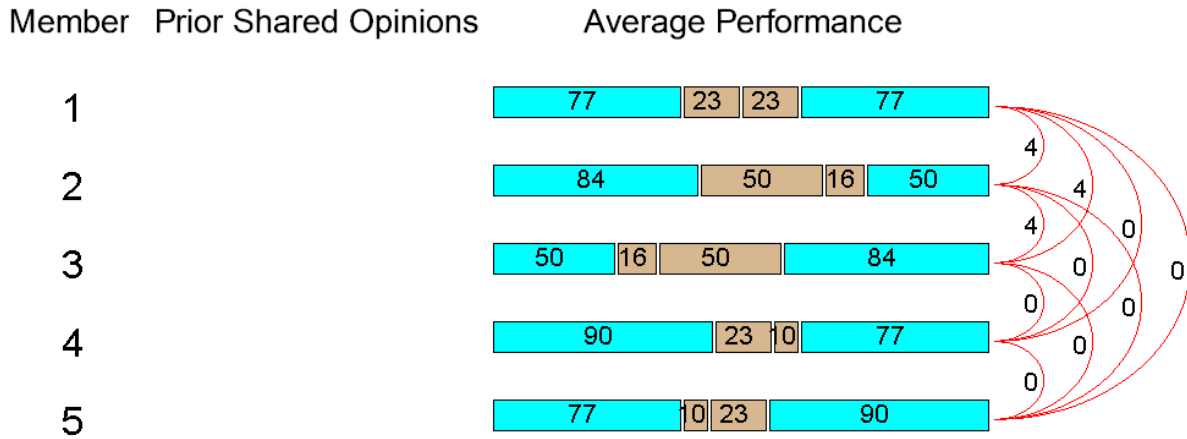
The group has not yet reached consensus. Please vote again:

Noise   Signal

Your estimate was

Your last (secret) vote was: Signal

Member   Prior Shared Opinions   Average Performance

1 | 77 | 23 | 23 | 77 |
                                                4
2 | 84 | 50 | 16 | 50 | 4
                                                4   0
3 | 50 | 16 | 50 | 84 | 0   0
                                            0   0
4 | 90 | 23 | 10 | 77 | 0
                                            0
5 | 77 | 10 | 23 | 90 |

Member 2 is now sharing a new opinion: Noise

The group has not yet reached consensus. Please vote again:

| Noise | | Signal |