Test of a Level of Aspiration model of group decision making: Non-obvious group preference reversal due to an irrelevant alternative

Ohtsubo, Yohsuke

Miller, Charles E.

(Citation)
Journal of Experimental Social Psychology,44(1):105-115

(Issue Date)
2008-01

(Resource Type)
journal article

(Version)
Accepted Manuscript

(URL)
https://hdl.handle.net/20.500.14094/90001045

Running head: LEVEL OF ASPIRATION MODEL

Test of a Level of Aspiration Model of Group Decision Making:

Non-Obvious Group Preference Reversal Due to an Irrelevant Alternative

Yohsuke Ohtsubo Charles E. Miller

Kobe University Northern Illinois University

Published in *Journal of Experimental Social Psychology*, Vol. 44, No. 1, (2008)

Abstract

We conducted two experiments that investigated a non-obvious prediction of a Level of Aspiration (LOA) model of group decision making. In both experiments, groups chose among three alternatives, A, B, and C, with C always being least preferred by all of the group members. In each experiment, by manipulating the nature of alternative C, we created two conditions differing in whether only A or only B was acceptable to all members. In support of the LOA model, groups tended to choose the alternative acceptable to all members, even when (Experiment 2) a majority of members most preferred a different alternative. Implications of the results for the LOA model are discussed in the context of social choice theory.

Keywords: Group Decision Making, Level of Aspiration Model, Preference Strength

Test of a Level of Aspiration Model of Group Decision Making: Non-Obvious Group Preference Reversal Due to an Irrelevant Alternative

Group decision making is ubiquitous in everyday life. A familiar example is when an academic department hires a new faculty member. Typically, a search committee is appointed to choose the applicant most suited to the department. In determining the most suitable applicant, each member of the search committee not only may rank order applicants in terms of their suitability, but may also find some applicants much more suitable than others. This example illustrates two aspects of preference: preference order and preference strength.

Intuitively, we expect group decisions to incorporate the group members' strength of preferences as well as their preference orders. For example, even if two members of a three-person search committee rank order a particular applicant relatively high, the committee may be unlikely to choose that applicant if the remaining member finds the applicant absolutely unsuitable. In the present article, we report two experimental tests of a non-obvious prediction of a group decision making model—the LOA model—that incorporates group members' preference strengths. In particular, suppose that a group is to choose one of three job applicants, A, B, or C, and that Applicant C is inferior to Applicants A and B, and thus has virtually no chance of being chosen. In effect, the group's task is to choose the better applicant between A and B. We demonstrate that, as predicted by the LOA model, the group's choice of A or B may depend on the specific nature of Applicant C, even though C is the unanimously least-preferred applicant.

The Paradox of Voting and Preference Strength

The effect of preference strength has been somewhat neglected in research on group decision making. Whether or how preference strength may be reflected in group decisions is,

however, not only an interesting empirical question, but also an important theoretical issue in light of the famous paradox of voting, the Condorcet paradox (Arrow, 1963; Black, 1958). Imagine a situation in which a group of three individuals (I, II, and III) must choose one of three alternatives (A, B, or C). Individual I prefers A to B, and B to C; Individual II prefers B to C, and C to A; and Individual III prefers C to A, and A to B. The Condorcet Paradox shows that if the alternatives are voted on pairwise using majority rule, the outcome depends on the order in which the pairwise votes are taken. If the first vote is between A and B, Alternative A will be chosen because a majority of the members (I and III) prefers A to B. Then, when A is voted on against the remaining alternative, C, Alternative C will be the group choice because a majority (II and III) prefers C to A. Is it possible from this result to conclude definitively that C is the group's most preferred alternative? Unfortunately, it is not. For example, if Alternative C is first voted on against B, C will be defeated and B will be chosen, because a majority of the members (I and II) prefers B to C. Then, when B is voted on against A, Alternative A will be chosen because it is preferred by a majority (I and III). In general, the paradox lies in the fact that the group's preference order is not transitive, viz., the group prefers A to B and B to C and C to A.

Social choice theorists have noted that in certain instances the Condorcet paradox might be resolved, or at least weakened, if the group decision were to take into account not only preference order, but also preference strength (Coleman, 1966). Suppose, for example, that Individual I very strongly prefers Alternative A over the other two alternatives, whereas the two remaining group members have very weak preference differences among the three alternatives. In this case, it might be quite reasonable for the group to choose Alternative A. Social choice theorists do not generally find such a solution compelling, however, because it involves interpersonal comparisons

of subjective utility (Arrow, 1963). Establishing objectively that Individual I's strength of preference is greater than the other two members' preference strengths is problematic. Nevertheless, in actual decision making groups, members' preference strengths may be revealed, at least to some extent, as they interact with each other (see Ohtsubo & Kameda, 1998, for a similar argument in the bilateral bargaining context).

Level of Aspiration Model of Group Decision Making

Harnett (1967) proposed a Level of Aspiration model (LOA model) of group decision making, and noted an interesting way in which group decisions might incorporate group members' preference strengths. Harnett's LOA model of group decision making is based on Siegel's (1957) LOA model of individual choice behavior. Siegel assumes that individual decision makers not only rank order alternatives, but also have a level of aspiration according to which some alternatives may be categorized as acceptable, and others as unacceptable. Siegel provides a concrete definition of the level of aspiration. Suppose that one finds Alternative A more desirable than Alternative B, and B more desirable than C. According to Siegel, one can further compare the relative differences in desirability between alternatives adjacent in the preference order, i.e., the difference between A and B and the difference between B and C. The level of aspiration is defined as lying between the pair associated with the largest difference. For example, if the difference in desirability between A and B is greater than the difference between B and C (e.g., if A is worth $10,000, B $9,000, and C $8,900), then the level of aspiration is said to be between A and B, and only A exceeds the level of aspiration and is regarded as acceptable. On the other hand, if the B-C desirability difference is greater than the A-B difference (e.g., if A is worth $10,000, B $9,000, and C $100), then the level of aspiration is between B and C, and both A and B are considered

acceptable. Thus, Siegel's model uses relative preference strength to define level of aspiration and categorize alternatives as acceptable or unacceptable.

Harnett applied Siegel's (1957) notion of level of aspiration to group decision making. Harnett's LOA model predicts that groups will choose the alternative that is acceptable to the greatest number of members. Thus, if Alternative A is acceptable to all the members of a group, whereas the other alternatives are acceptable to fewer members, the LOA model would predict A as the group decision.

The LOA model has been tested empirically by Castore, Peterson, and Goodrich (1971), by Lieberman (1971), and by Harnett himself (1967). Castore et al. tested the ability of the model to account for group polarization or group choice shifts. Their experiment studied 3- and 4-person groups, involved face-to-face discussion, used Choice Dilemma types of items, and measured LOA via the method outlined by Siegel (1957). The model correctly predicted 80% of all group decisions. Moreover, although the model did not account for the exact magnitude of every group choice shift, the direction of the model's predictions (risky or cautious shift or no shift) was correct in all instances (see Miller, 1989).

Lieberman (1971) tested the LOA model in the context of the Condorcet paradox. Group members' preference orders were manipulated through the use of monetary rewards, and their levels of aspiration were assessed via a questionnaire (details of which are not provided in Lieberman's article.) After assessing the group members' levels of aspiration, Lieberman had the members engage in sequential exchanges of written messages to achieve unanimous agreement. The measured levels of aspiration were used to predict the group decisions. The LOA model sometimes failed to predict a specific alternative (when two or more alternatives are tied as

acceptable to the largest number of members, the LOA model's prediction is indefinite), but the model was well-supported in those instances in which it did make a specific prediction.

Harnett's (1967) own experiment also supported the LOA model. Harnett manipulated the preference orders of three alternatives and measured level of aspiration using Siegel's (1957) method. Experimental groups made their decisions via sequential bidding until unanimous agreement was achieved. Unlike Lieberman's experimental groups, in which no majority factions were present, Harnett's experimental groups included a majority faction. It is well-known that the preference of an initial majority faction often becomes the group decision, consistent with the majority-wins model of group decision making (Davis, 1973). Using the assessed levels of aspiration, Harnett composed groups so that the LOA model's prediction would be contrary to the prediction of the majority-wins model: A majority of group members preferred an alternative that was unacceptable to a minority member, whereas the minority member's most preferred alternative was acceptable to all of the group members (cf. Experiment 2 of this article). Hartnett's results generally supported the predictions of the LOA model. It might be supposed that the failure of the majority-wins model was due to the requirement for unanimous agreement. It is worth noting in this regard, however, that group decision making experiments have often found that groups tend to choose an alternative initially preferred by a majority, even if there is a formal requirement of unanimity (see Miller, 1989, for a review).

In summary, although there have been only a few experimental tests of the LOA model, the results of these tests are promising, and further testing appears warranted. As detailed below, the primary purpose of the present experiments was to test a non-obvious effect predicted by the LOA model, and thus to provide a test of the model different from the tests of previous studies.

Overview of the Two Experiments

Both of our experiments employed a preference manipulation involving a variant of the hidden profile paradigm (Stasser & Titus, 1985). We shall first explain the experimental task and level of aspiration manipulation common to both experiments, and then explain how the manipulation leads to a non-obvious prediction.

Group Decision Making Task. In the variant of the hidden profile paradigm that we employed, each group member had a different set of information. Each information set contained both shared items and unshared items. The shared items were assigned to all three or all four members of a group, whereas the unshared items were assigned to one or two members. The decision task was a hypothetical hiring decision in which participants were instructed to imagine that they were members of the personnel staff in a Japanese branch of a foreign company. Participants were informed that the company's employment policy emphasized two criteria: applicants' English skills, assessed by scores on the TOEIC (Test of English for International Communication), and applicants' interpersonal skills, assessed by a job interview. Participants were asked to choose the most desirable of three applicants, A, B, and C, after reviewing their profiles.

The three applicants' profiles were such that the TOEIC score favored Applicant A, whereas the interview score favored Applicant B. The third applicant, Applicant C, was inferior to the other two applicants in terms of both the TOEIC score and the interview score (see Table 1). Thus, although there were three applicants to decide among, groups in reality chose between the two applicants who were stronger with regard to the two important criteria. Given such a pattern of applicant profiles, participants' preference orders were manipulated by unevenly distributing

the profile items among group members. In particular, prior to group discussion, some participants were provided with the three applicants' TOEIC scores but not the applicants' interview scores. Other participants were provided with the three applicants' interview scores but not the applicants' TOEIC scores (again, see Table 1). It was anticipated that the participants informed of the TOEIC scores would prefer Applicant A to Applicant B, and Applicant B to Applicant C, whereas the participants informed of the interview scores would prefer Applicant B to Applicant A, and Applicant A to Applicant C.

Participants' levels of aspiration were manipulated based on Siegel's (1957) definition. Suppose that the three applicants' TOEIC scores were 750, 600, and 230 for A, B, and C, respectively (A, B, and C1 in Table 1). It is plausible that such scores would lead one to find Applicant A more desirable than Applicant B, and B more desirable than C. It is also plausible that one would find the B-C desirability difference greater than the A-B difference. Therefore, given such TOEIC scores, both A and B would exceed the LOA, and both would be acceptable. On the other hand, suppose that the three applicants' TOEIC scores were 750, 600, and 580 (A, B, and C2 in Table 1). In this case, the A-B desirability difference would be greater than the B-C difference, and only A would exceed the LOA and be acceptable. (See Ohtsubo & Watanabe, 2003, for evidence that supports manipulating level of aspiration in this manner.) Notice that a change in the weakest applicant's profile (i.e., C's profile) is sufficient to manipulate of the level of aspiration.

Purposes and Hypotheses of the Two Experiments. In Experiment 1, four-person groups engaged in a hypothetical hiring decision task. Two of the four group members (Members I and II) were provided with the three applicants' TOEIC scores, and hence tended to prefer Applicant A to Applicant B, and Applicant B to Applicant C (hereafter, "Applicant C" means either Applicant C1

or C2). The other two members (III and IV) were provided with the three applicants' interview scores, and hence tended to prefer Applicant B to Applicant A, and Applicant A to Applicant C (see the upper half of Table 1).

The group members' levels of aspiration were manipulated to create two experimental conditions. In one condition (the C1 condition), the TOEIC scores for Applicants A, B, and C1 were 750, 600, and 230, respectively, whereas the interview scores were 75, 95, and 70, respectively. These scores were such that Members I and II, who were informed of the TOEIC scores, would find both A and B acceptable, whereas Members III and IV, who were informed of the interview scores, would find only B acceptable. Thus, in the C1 condition, the LOA model predicted that groups would tend to choose Applicant B, who was acceptable to all four members.

In the other condition (the C2 condition), the TOEIC scores were 750, 600, and 580 for Applicants A, B, and C2, respectively, whereas the interview scores were 75, 95, and 35, respectively. These scores were such that Members I and II, who were informed of the TOEIC scores, would find only A acceptable, whereas Members III and IV, who were informed of the interview scores, would find both B and A acceptable. Thus, in the C2 condition, the LOA model predicted that groups would tend to choose Applicant A, who was acceptable to all four members.

The prediction of the LOA model that Applicant B would be relatively more likely, and Applicant A relatively less likely, to be the group choice in the C1 condition than in the C2 condition represents a non-obvious preference reversal at the group level. The only difference between the C1 and C2 conditions is in the profile of scores for Applicant C, who is the least-preferred candidate of all the group members in both conditions. The score profiles for the two more preferred applicants, A and B, are exactly the same in the C1 condition as in the C2

condition. Nevertheless, the LOA model predicts a reversal in group preference for Applicants A and B between the C1 and C2 conditions.

Experiment 2 tested this same non-obvious group preference reversal prediction, but also went beyond Experiment 1 in the following way: Note that in Experiment 1, there was no majority faction in the groups. Experiment 1 was therefore a weak test of the LOA model in the sense that the majority-wins model (Davis, 1973) was not applicable. Experiment 2 introduced the presence of a majority faction and tested the rather more straightforward prediction of the majority-wins model against the non-obvious prediction of the LOA model.

In Experiment 2, three-person, rather than four-person, groups engaged in the hypothetical hiring decision task. One of the three members (Member I) was informed of the applicants' TOEIC scores, and hence tended to prefer Applicant A to Applicant B, and Applicant B to Applicant C. The other two members (II and III) were informed of the applicants' interview scores, and hence tended to prefer B to A, and A to C (see the lower half of Table 1). Members II and III constituted a majority faction in Experiment 2. The level of aspiration manipulations in Experiment 2 were comparable to those in Experiment 1. The manipulations were such that, in the C1 condition, Member I would find Applicants A and B acceptable, whereas Members II and III would find only Applicant B acceptable. In the C2 condition, Member I would find only A acceptable, whereas Members II and III would find B and A acceptable. Therefore, the LOA model predicted a preference reversal, with groups being relatively more likely to choose Applicant B in the C1 condition, but Applicant A in the C2 condition. Moreover, in the C2 condition, the prediction of the LOA model was for the group decision favored initially by a minority (i.e., Member I), and thus was contrary to the prediction of the majority-wins model.

To summarize, the purpose of both Experiment 1 and Experiment 2 was to test the non-obvious preference reversal prediction of the LOA model. Beyond that, the purpose of Experiment 2 was also to test the LOA model against the majority-wins model.

Experiment 1

Method

Participants, Design, and Materials. Participants were 188 undergraduates (119 males and 62 females, 7 unidentified) enrolled in an organizational psychology class at Nara University. Participation in the study partially fulfilled a course requirement. Participants took part in groups consisting of four persons, for a total of 47 groups. Participants were randomly assigned to experimental condition (C1 or C2) and to member type (I, II, III, or IV) within each group.

Each participant received partial profiles of the three applicants. The partial profiles were summarized in a resume format. Each applicant's resume contained both the experimentally manipulated items and filler items, accompanied by the applicant's name, date of birth, place of birth, and sex (all applicants were described as male). The filler items consisted of computer skill, bookkeeping skill, driver's license status, knowledge test score, volunteer experience, and education. The filler items involve information commonly included in Japanese job applicants' resumes. The scores of the filler items were arranged so that they would not be inconsistent with the expected preference order as manipulated by the TOEIC scores or the interview scores (see Table 1).

Procedure. Participants were told by the experimenter that their task was to determine the best of three applicants for a job that emphasized applicants' English fluency, assessed by the TOEIC, and interpersonal skills, assessed by a job interview. The experimenter explained that

each group member had only portions of the full set of profiles. Each participant was given a booklet containing the applicants' profiles, and asked to evaluate the applicants on a scale from 0 (very undesirable) to100 (very desirable). Participants were not asked any questions regarding their level of aspiration, so as to avoid the suggestion that level of aspiration be used in the group decision making to follow. After all group members had completed the evaluation of the three applicants, they were given one extra minute to memorize the applicants' profiles. After one minute had elapsed, the experimenter collected the booklets and gave out a group decision sheet. Each group was asked to discuss the applicants and achieve unanimous agreement about which applicant was the best. When a group had made its decision, the experimenter collected the group decision sheet and gave the group an evaluation sheet asking the group to evaluate the three applicants on a scale from 0 (very undesirable) to 100 (very desirable). After completing the group evaluation, members were asked to indicate individually which applicant they personally thought was best, and to rate the three applicants from 0 (very undesirable) to 100 (very desirable). Once all the experimental sessions had been completed, a debriefing session was conducted in class by the first author.

Results and Discussion

Manipulation Checks. We assessed group members' preference orders prior to group discussion by asking them to evaluate the desirability of the three applicants. We anticipated that Members I and II would each prefer Applicant A to Applicant B, and Members III and IV would each prefer Applicant B to Applicant A. The left-hand columns of Table 2 show the mean desirability ratings of the three applicants as a function of condition (C1 vs. C2) and member type (I-II vs. III-IV). (To compute the means, we averaged the ratings of Members I and II and the

ratings of Members III and IV within each group, so that for each group there were two data points representing the two factions.) Examination of the desirability ratings showed that there were only two participants whose ratings were not consistent with the expected preference orders. The two groups to which these participants belonged were omitted from the data set. One other group was also omitted because it failed to provide its group decision in the instructed manner, and its decision was not readily interpretable. Thus, there were 44 groups retained in the data set (22 in each experimental condition), all of which had preference orders that were consistent with expectations.

Hypothesis Testing. The experimental hypothesis predicted that groups in the C1 condition would be relatively more likely to choose Applicant B, and relatively less likely to choose Applicant A, than would groups in the C2 condition. The frequencies of groups that chose Applicants A and B, respectively, in each condition are shown in the upper half of Table 3. As predicted, groups in the C1 condition tended to choose Applicant B (19 of 22 groups), whereas groups in the C2 condition tended to choose Applicant A (13 of 22 groups), $\chi^2(1, N = 44) = 9.82$, $p = .002$. Because there was one cell with fewer than five observations, Fisher's exact test was also performed, and it too revealed that the probability of obtaining the observed pattern of data by chance was .002 (one-sided). Therefore, the hypothesis was supported.

Groups also rated the desirability of each applicant (from 0 to 100). Group evaluations were expected to be consistent with the group decisions: Applicant B would tend to be rated better in the C1 condition, whereas Applicant A would tend to be rated better in the C2 condition. One group in the C2 condition apparently misunderstood the rating scale as 0 to 10. Data points from that group were considered as outliers (*Z*-scores were -5.05 for Applicant A, and -5.08 for

Applicant B), and the group was omitted from this analysis (i.e., $n = 21$ for the C2 condition). The mean desirability of Applicant A was higher in the C2 condition (M =81.90, SD = 6.89) than in the C1 condition (M = 77.09, SD = 10.45), whereas the mean desirability of Applicant B was higher in the C1 condition (M = 88.73, SD = 7.51) than in the C2 condition (M =79.10, SD = 9.66). A 2 (Condition: C1 vs. C2) x 2 (Applicant: A vs. B) ANOVA, with the latter factor as repeated measures, revealed a significant main effect of applicant, $F(1, 41) = 6.41$, $p = .015$, $\eta^2 = .135$, and a significant Condition x Applicant interaction, $F(1, 41) = 17.16$, $p < .001$, $\eta^2 = .295$. The interaction effect supported the experimental hypothesis. The unexpected main effect of applicant reflects a tendency for groups to find Applicant B more desirable than Applicant A. This effect, although not intended, is consistent with the pattern of group decisions: Collapsing across the two conditions, Applicant B was chosen by 28 of 44 groups ($p = .048$ by the sign test, assuming that the two alternatives are equally likely to be chosen).[1]

Post-Discussion Individual Preferences. After completing the group choice and evaluation session, each group member was asked individually to indicate the applicant he or she most preferred, and to rate the three applicants' desirability again. We compared the group's choice and each participant's individual post-discussion choice to determine whether there were any holdout members, i.e., members whose post-discussion choice was different from the group's choice. Although there were 23 of the 44 groups in which at least one holdout member existed after the group decision had been made, the overall holdout rate was a modest .15. We then analyzed the post-discussion ratings of Applicants A and B, shown in the right-hand columns of Table 2. A 2 (Condition: C1 vs. C2) x 2 (Applicant: A vs. B) x 2 (Member type: Members I & II vs. Members III & IV) ANOVA, with the latter two factors as repeated measures, revealed a

significant main effect of condition, $F(1, 42) = 4.48$, $p = .04$, $\eta^2 = .096$, a significant main effect of applicant, $F(1, 42) = 11.43$, $p = .002$, $\eta^2 = .214$, a significant Condition x Applicant interaction, $F(1, 42) = 15.05$, $p < .001$, $\eta^2 = .264$, and a significant Applicant x Member type interaction, $F(1, 42) = 13.44$, $p = .001$, $\eta^2 = .242$. The significant main effect of applicant is due to the fact that Applicant B was more attractive than Applicant A. The significant interaction effect involving the member type factor appears to be a residue of the initial manipulation. The most pertinent effect, the significant Condition x Applicant interaction, is consistent with the group level preference reversal effect.

Discussion. Experiment 1 tested the LOA model in a situation where no majority faction was present (i.e., two alternatives were tied as most-preferred). In the absence of a majority faction, groups exhibited the hypothesized preference reversal in their choices, and were likely to choose the alternative acceptable to the largest number of members. Recall that Harnett's (1967) original study showed that an alternative that was initially preferred by a minority of group members was chosen by the group when it was acceptable to all of the group members. Therefore, it was anticipated that the preference reversal observed in Experiment 1 would also be observed in a setting where majority and minority factions were present. In Experiment 2, we included a condition in which the LOA model would predict the choice of an alternative preferred by a minority, rather than by a majority, of group members.

Experiment 2

Method

Pilot Test. In the version of materials used in Experiment 1, Applicant B was relatively more attractive than intended. In an attempt to eliminate, or at least reduce, such an effect in

Experiment 2, we created two versions of materials and conducted a pilot test. Ninety-one participants drawn from a population comparable to that of the main study participated in the pilot study. Each participant was randomly assigned to one cell of a 2 (Version: 1 vs. 2) x 2 (Condition: C1 vs. C2) factorial design. Participants were given a complete set of profiles of the three applicants corresponding to their condition. Participants indicated their choice of the best applicant and then evaluated the three applicants on a scale from 0 (very undesirable) to 100 (very desirable). The results revealed that for Version 1 there was a marginally significant tendency to choose different applicants in the different conditions. Therefore, we decided not to use this version for Experiment 2.

Version 2 of the three applicant profiles is summarized in the lower half of Table 1. For Version 2, Applicant A was chosen by 10 of 22 participants in the C1 condition, and by 9 of 23 participants in the C2 condition, $\chi^2(1, N = 45) = .67$, *ns*. Although Applicant B tended to be chosen more often than Applicant A (i.e., across the two conditions 26 of 45 participants chose B), this difference was not statistically significant by the sign test. In addition, applicants' desirability scores did not differ between the two conditions: for Applicant A, $M = 83.14$ in the C1 condition and $M = 83.57$ in the C2 condition, $t(43) = .14$, *ns* (two-tailed), $d = .04$; for Applicant B, $M = 84.00$ in the C1 condition and $M = 85.04$ in the C2 condition, $t(43) = .41$, *ns* (two-tailed), $d = .12$; for Applicant C, $M = 51.00$ in the C1 condition and $M = 48.87$ in the C2 condition, $t(43) = .38$, *ns* (two-tailed), $d = .11$. Moreover a 2 (Condition: C1 vs. C2) x 2 (Applicant: A vs. B) ANOVA, with the latter factor as repeated measures, revealed neither significant main effects nor a significant interaction effect. Therefore, in this version of the profiles, participants tended to find Applicants A and B about equally preferable across conditions.[2] Based on these pilot test results, we decided

to use Version 2 of the profiles in Experiment 2. Note that this pilot test also serves the same purpose as the individual control condition in Experiment 1 (see footnote 1).

Participants, Design, Materials, and Procedure. Participants in the main study were 117 undergraduates (72 males and 45 females) enrolled in an experimental psychology class or an introductory psychology class at Nara University. Participation in the study partially fulfilled a course requirement. Each group consisted of three persons, for a total of 39 groups in the experiment. There were two experimental conditions comparable to those of Experiment 1. Participants were randomly assigned to one of the two conditions (C1 or C2) and to one of the three member types (I, II, or III) within each group. The experimental procedure was the same as the procedure used in Experiment 1, except that the group discussions were video recorded in Experiment 2. We do not report the video data, however, because they did not reveal any significant results relevant to present purposes.

Results and Discussion

Manipulation Checks. As in Experiment 1, we examined the pre-discussion desirability ratings to determine whether the preference orders were as intended. In each condition, there were four groups in which at least one member did not have the expected preference order. These groups were omitted from the data set. Also, one group failed to reach consensus within the required 20 minutes, and there was no indication that it would reach a decision soon. The experimenter terminated this group's discussion. Thus, 30 groups (15 in each condition) were retained for the data analyses. All of the retained groups' preference structures met the expected structure, viz., Member I preferred A to B, and B to C, whereas Members II and III preferred B to A, and A to C. The left-hand column of Table 4 shows the mean desirability ratings given to the

applicants as a function of member type and condition.

Hypothesis Testing. As predicted in the experimental hypothesis, groups in the C1 condition were relatively more likely to choose Applicant B (13 of 15 groups), and groups in the C2 condition were relatively more likely to choose Applicant A (9 of 15 groups), $\chi^2(1, N = 30) = 7.03$, $p = .008$ (see the bottom half of Table 3). Because there was one cell with fewer than five observations, Fisher's exact test was also performed, and it revealed that the probability of obtaining the observed data pattern by chance was .01 (one-sided). The experimental hypothesis was therefore supported.

Groups also rated the desirability of each applicant from 0 to 100. Consistent with the group decision pattern, the mean desirability of Applicant A was higher in the C2 condition ($M = 84.80$, $SD = 4.28$) than in the C1 condition ($M = 78.07$, $SD = 11.22$), whereas the mean desirability of Applicant B was higher in the C1 condition ($M = 87.87$, $SD = 5.83$) than in the C2 condition ($M = 81.20$, $SD = 7.86$). A 2 (Condition: C1 vs. C2) x 2 (Applicant: A vs. B) ANOVA, with the latter factor as repeated measures, revealed a significant interaction effect, $F(1, 28) = 18.29$, $p < .001$, $\eta^2 = .395$, in support of the hypothesis.

Post-Discussion Individual Preferences. After completing the group choice and evaluation session, each member was asked again to indicate his or her most-preferred applicant, and to rate the three applicants' desirability. As in Experiment 1, we examined whether there were any holdout members. There were no holdouts in 22 of the 30 groups, and the holdout rate was a relatively low .12. As in Experiment 1, we then analyzed the post-discussion ratings of Applicants A and B, shown in the right-hand columns of Table 4. A 2 (Condition: C1 vs. C2) x 2 (Applicant: A vs. B) x 2 (Member type: Member I vs. Members II & III) ANOVA, with the latter two factors as

repeated measures, revealed a significant Condition x Applicant interaction, $F(1, 28) = 13.27$, $p = .001$, $\eta^2 = .322$, and a significant Applicant x Member type interaction, $F(1, 28) = 19.45$, $p < .001$, $\eta^2 = .410$. The latter significant interaction effect may be a residue of the initial manipulation. The former effect is consistent with the preference reversal at the group level.

Discussion. The results of Experiment 2 replicated the results of Experiment 1 with regard to the non-obvious preference reversal effect in group choice predicted by the LOA model. Experiment 2 also tested the LOA model against the majority-wins model. In particular, the preference configurations in the C2 condition of the experiment were designed so that the LOA model predicted a minority-wins group decision. The results confirmed the LOA model's prediction. It is worth noting, however, that 6 of the 15 groups in the C2 condition chose the majority-preferred alternative. We shall return to the issue of the LOA model's predictive accuracy shortly.

General Discussion

Both Experiments 1 and 2 manipulated participants' preference strength, and thereby level of aspiration, while holding preference order constant. In Experiment 1, four-person groups consisting of two members who preferred Applicant A to Applicant B and two members who preferred Applicant B to Applicant A engaged in the group decision making task. Members' preference strengths and levels of aspiration were manipulated by systematically varying the nature of a third, least-preferred applicant, C1 or C2. As predicted by the LOA model, groups tended to choose the applicant who was acceptable to (i.e., exceeded the levels of aspiration of) all four members, thus resulting in a between-conditions group preference reversal with regard to Applicants A and B. In Experiment 2, the LOA model was tested using three-person groups in

which, in the C2 condition, the model made a prediction contrary to the majority-wins model. As predicted by the LOA model, groups tended to choose the applicant who was acceptable to all three members—to some extent even when that applicant was initially most preferred by only one member—and the preference reversal was again demonstrated.

Irrationality of Group Decision Making

The paradox of voting is an example of irrationality associated with the use of majority vote in group decision making. Group preferences derived through pairwise majority voting may be intransitive, and thus the eventual group decision may depend on the order of voting upon alternatives. An immediate implication of this is that group decisions can be manipulated. For example, a chairperson who has the prerogative of setting a group's agenda might be able to manipulate the group's decisions by scheduling the order in which alternatives are brought to vote. The ostensibly democratic use of majority vote may be quite autocratic in actuality.

As we noted earlier, some social choice theorists have suggested that the paradox of voting might be resolved or weakened if group decisions were to take into account not just preference order, but preference strength. The LOA model implies that decisions made by group discussion incorporate not only group members' preference orders, but also the strength of their preferences. Thus, group discussion tends to yield decisions that reflect the members' preferences more fully than does voting, and in this sense group discussion may produce better decisions—group preferences that are free of intransitivities and thus less susceptible to manipulation.

Unfortunately, the LOA model and the results of our experiments imply that even if decision making by group discussion is not subject to preference intransitivity, it is subject to a different sort of irrationality. In particular, the group preference reversal observed in our

experiments indicates that decision making via group discussion violates the independence of irrelevant alternatives criterion, proposed by Arrow (1963) as a requirement for the rationality of social choice. As already detailed, in both of our experiments the profiles of Applicants A and B were held constant across experimental conditions, with the difference between conditions residing solely in the nature of the least preferred alternative, C1 or C2. Nevertheless, the frequency with which groups chose A versus B differed from one condition to the other. Thus, the group level preference order for the two alternatives, A and B, is attributable to the nature of the unrelated, and thus irrelevant, third alternative, C1 or C2.

The fact that decision making via group discussion may violate the independence of irrelevant alternatives condition means that such decision making may be manipulated. For example, a manipulative chairperson might eliminate some apparently unpopular alternatives before beginning group discussion and decision making. This might be justified by claiming that the group does not need to consider alternatives that are not endorsed by anyone and that the group is unlikely to choose. This seemingly innocuous trimming of unpopular alternatives prior to discussion, however, might change the group decision by shifting the level of aspiration of one or more group members. Once again, then, an apparently democratic decision process might actually be rather autocratic.

Of course, the external validity of our results has yet to be tested.[3] It could be argued that when groups discuss real issues in which the group members are deeply involved, the strength of members' preferences might be more stable than in the case of experimentally manipulated preferences. Moreover, only certain configurations of member preferences (e.g., those in our experiments) would provide opportunity for manipulation, and a chairperson would need

considerable information about those preferences to take advantage of such an opportunity. Therefore, group decision making via discussion might be less vulnerable to manipulation in real world group contexts than in experimental settings. Nevertheless, preference strength, at least as measured in terms of the level of aspiration, is generally less stable than preference order, and one might question whether it is desirable to make group decisions in a manner that is potentially vulnerable to such instability (see Ohtsubo & Watanabe, 2003, for a similar argument in the approval voting context). These issues are related to a developing interest in the procedural justice of group decision making (see Kameda, 1996, for a review), and deserve further empirical investigation.

Predictive Accuracy of the LOA Model

Although the LOA model's prediction of a group preference reversal was supported in both Experiments 1 and 2, there was one condition in each experiment for which the model correctly predicted only about 60% of the group decisions. In the C2 Condition of Experiment 2, for example, in which the LOA model was pitted against the majority-wins model, 60% of the group decisions were as predicted by the LOA model (see Table 3). Although we find it impressive that 60% of the groups did *not* choose the alternative preferred by a majority, this result is obviously inconclusive with respect to the relative accuracy of the LOA and majority-wins models. And of course, the fact that 40% of the groups in this condition made a choice contrary to that predicted by the LOA model raises the question of how the predictive accuracy of the model could be improved.

One way of possibility might be to consider how level of aspiration is defined and measured. Although we adopted the Harnett-Siegel definition of level of aspiration for our

manipulations of level of aspiration, the Harnett- Siegel definition of LOA does not allow for the possibility of an individual finding all (or none) of a set of alternatives acceptable. That is, defining level of aspiration in terms of the largest preference difference between adjacent alternatives in a preference ordering necessarily means that not all of the alternatives can fall above (or below) the LOA. However, defining LOA to include the possibility that all (or no) alternatives are acceptable seems unobjectionable on conceptual grounds, and might improve the accuracy of the LOA model. Another possibility might be to define LOA as more or less gradual or as involving more than two categories of alternatives (i.e., highly acceptable vs. acceptable vs. unacceptable). We suggest that future research on the LOA model explore different definitions of the LOA construct, different methods of measuring the construct, and the relationship of these different definitions and measures to the model's predictive accuracy.

Applicability of the LOA model

In addition to the predictive accuracy of the LOA model, the applicability of the model should be considered. There are various group patterns of preferences and preference strengths for which the LOA model and the majority-wins model predict the same outcome. For example, the two models make the same prediction in instances where all group members have levels of aspiration that are exceeded by only their most-preferred alternative. Such instances might be especially likely to arise in situations that involve only a small number of alternatives (e.g., jury decision making in criminal cases where possible verdicts are guilty, not guilty, or guilty of a lesser charge). In these instances, the LOA model would have no predictive advantage over the majority-wins model, and might be regarded as inferior on the basis of parsimony, because it requires more information about group members' preferences than does the majority-wins model.

On the other hand, when there are many alternatives, the LOA model might be applied more readily than the majority-wins model. Suppose that a group has to choose one alternative from among, say, eight or nine alternatives. In such a case, different members may often have different alternatives as their first preference, and the majority-wins model may be incapable of providing a meaningful prediction unless it is combined with some sort of subrule (e.g., plurality rule) or some type of multi-stage procedure. The LOA model, in contrast, might provide a meaningful prediction, based on preference strength and members' levels of aspiration. This suggests that the LOA model may have a different niche as a predictive model than does the majority-wins model: The LOA model may apply more to tasks involving a relatively large number of alternatives, whereas the majority-wins model may apply more to tasks involving a relatively small number of alternatives.

Conclusion

The two experiments demonstrated that decisions via group discussion can be unstable due to instability in individual preference strengths. In particular, the two experiments showed that a group's preference between two alternatives could be reversed due to the relative strength of preference for these alternatives against a third, undesirable alternative. Harnett's (1965) LOA model provided a conceptual framework for testing this possibility, although the definition of the level of aspiration might be improved. As we have pointed out, issues of the external validity and applicability (or boundary conditions) of the LOA model await further investigation. Whether the sort of preference reversal observed in the present study is wide-spread in real-world decision making may be an important consideration in any debate about the defensibility of group discussion as a procedure for making decisions.

References

Arrow, K. J. (1963). *Social choice and individual values* (2nd ed.). New York: Wiley.

Black, D. (1958). *The theory of committees and elections*. Cambridge, England: Cambridge University Press.

Castore, C. H., Peterson, K., & Goodrich, T. A. (1971). Risky shift: Social value or social choice? *Journal of Personality and Social Psychology, 20,* 487-494.

Coleman, J. S. (1966). Foundations for a theory of collective decisions. *American Journal of Sociology*, *71*, 615-627.

Davis, J. H. (1973). Group decision and social interaction: A theory of social decision scheme. *Psychological Review*, *80*, 87-125.

Harnett, D. L. (1967). A level of aspiration model for group decision making. *Journal of Personality and Social Psychology*, *5*, 58-66.

Kameda, T. (1991). Procedural influence in small-group decision making: Deliberation style and assigned decision rule. *Journal of Personality and Social Psychology*, *61*, 245-256.

Kameda, T. (1996). Procedural influence in consensus formation: Evaluating group decision making from a social choice perspective. In E. H. Witte & J. H. Davis (Eds.), *Understanding group behavior: Consensual action by small groups* (Vol. 1, pp. 137-161). Mahwah, NJ: Erlbaum.

Kameda, T., Ohtsubo, Y., & Takezawa, M. (1997). Centrality in sociocognitive networks and social influence: An illustration in a group decision making context. *Journal of Personality and Social Psychology*, *73*, 296-309.

Lieberman, B. (1971). Resolving the Condorcet paradox. *American Journal of Sociology*, *77*,

540-559.

Miller, C. E. (1989). The social psychological effects of group decision rules. In P. B. Paulus (Ed.), *Psychology of group influence* (2nd ed., pp. 327-355). Hillsdale, NJ: Erlbaum.

Ohtsubo, Y. (2005). Should information be redundantly distributed among group members? Effective use of group memory in collaborative problem solving. *Applied Cognitive Psychology*, *19*, 1219-1233.

Ohtsubo, Y., & Kameda, T. (1998). The function of equality heuristic in distributive bargaining: Negotiated allocation of costs and benefits in a demand revelation context. *Journal of Experimental Social Psychology*, *34*, 90-108.

Ohtsubo, Y., Masuchi, A., & Nakanishi, D. (2002). Majority influence process in group judgment: Test of the Social Judgment Scheme model in a group polarization context. *Group Processes and Intergroup Relations*, *5*, 249-261.

Ohtsubo, Y., & Watanabe, Y. (2003). Contrast effects and approval voting: An illustration of a systematic violation of the independence of irrelevant alternatives condition. *Political Psychology*, *24*, 549-559.

Siegel, S. (1957). Level of aspiration and decision making. *Psychological Review*, *64*, 253-262.

Stasser, G., & Titus, W. (1985). Pooling of unshared information in group decision making: Biased information sampling during discussion. *Journal of Personality and Social Psychology*, *48*, 1467-1478.

Tindale, R. S., & Sheffey, S. (2002). Shared information, cognitive load, and group memory. *Group Processes and Intergroup Relations*, *5*, 5-18.

Author Note

Yohsuke Ohtsubo, Department of Psychology, Faculty of Sociology, Nara University.

Charles E. Miller, Department of Psychology, Northern Illinois University.

This research was supported by the Japan Society of the Promotion of Science (Grant 14710062). We are grateful to Mami Yamakawa who contributed to devising the experimental materials. We also thank Yoshiko Ohmori and Yoko Ohno for their help in conducing Experiment 2. Correspondence concerning this article should be addressed to Yohsuke Ohtsubo, Department of Psychology, Faculty of Sociology, Nara University, 1500 Misasagi-chgo, Nara, 631-8502, Japan. E-mail: ohtsubo@daibutsu.nara-u.ac.jp

Footnotes

[1]Although the group data supported the hypothesis, this result alone does not assure that the preference reversal related to the profiles of the least preferred applicant (i.e., C1/C2) is uniquely a group-level phenomenon. In this study, unlike typical hidden profile studies, the cognitive load on each member was relatively low (especially if only the TOEIC and interview scores are considered relevant items), and it is not unreasonable to expect that group members might have pooled all important information through discussion (see Ohtsubo, 2005, and Tindale & Sheffey, 2002, for effects of cognitive load in group memory tasks). Thus, the observed preference reversal might have occurred at the individual level. To exclude this possibility, an additional 54 undergraduates (39 males, 15 females) from a population comparable to that of the main study took part in a control condition to assess individual choice. These participants were given the complete set of profiles of either Applicants A, B, and C1 or Applicants A, B, and C2. Participants' preferences for the applicants were assessed only once in this control condition. Exactly the same proportion of participants chose Applicant A in the C1 and C2 conditions (i.e., 10 of 27 in each). Therefore, no preference reversal was observed in individual choice behavior. Parenthetically, as in the main group decision making study, Applicant B was significantly more popular than Applicant A (p = .04 with the sign test).

[2]The materials employed in Experiment 2 differed in one respect from the materials used in Experiment 1. In Experiment 1, the filler items given to each participant were consistent with the primary items (i.e., with the three applicants' TOEIC scores or interview scores). That is, no participants received filler items that favored an applicant different from the one favored by the primary items. Because, in Experiment 2, we wished to examine the relationship between level of

aspiration and minority influence, we attempted to minimize a factor that is known to be related to the influence of minority members, i.e., social sharedness (Kameda, Ohtsubo, & Takezawa, 1997). Each group member in Experiment 2 was assigned one item that was unshared and two items that were shared with one other group member (see the lower half of Table 1). Therefore, the level of social sharedness was the same for all three members (i.e., the number of items shared with other members was constant across the three members). As a byproduct of this control for social sharedness, Member II was assigned one filler item (i.e., knowledge test scores) which favored Applicant A over Applicant B, but was assigned interview scores as a primary item, which favored Applicant B over Applicant A. Perhaps for this reason, a slightly greater proportion of participants in Experiment 2 than in Experiment 1 reported preference orders that did not meet our expectation (see the manipulation checks).

[3]With regard to external validity, one might ask whether some cultural difference is responsible for the reported results, given that both experiments were conducted in Japan. We admit that this is an open question. However, the results of Japanese group decision making experiments suggest some important similarities in behavior between groups of Japanese undergraduates and groups of U.S. undergraduates. For example, the majority-wins model provides a good fit for both Japanese and U.S. experimental groups engaging in judgmental tasks (e.g., Kameda, 1991; Kameda et al., 1997; Ohtsubo, Masuchi, & Nakanishi, 2002).

Table 1

Distribution of Applicant Profile Information Among Group Members

	Applicants				Members Who Were Assigned the Corresponding Item			
	A	B	C1	C2				
Experiment 1								
TOEIC	**750**	**600**	**230**	**580**	**I**	**II**		
Interview Score	**75**	**95**	**70**	**35**			**III**	**IV**
Filler Item:								
Computer Skill*	2nd	3rd	5th	3rd	I			
Bookkeeping Skill	2nd	3rd	N/A	3rd		II		
Drivers License	Yes	Yes	Yes	Yes	I	II		
Knowledge Test Score**	75	85	75	20			III	
Volunteer Experience	1 yr	2 yrs	1 yr	N/A				IV
Education***	UG	UG	UG	UG			III	IV

Table 1 *continued*

Table 1 *continued*

Experiment 2

TOEIC	**750**	**600**	**230**	**580**	**I**		
Interview Score	**75**	**95**	**70**	**35**		**II**	**III**
Filler Items							
Computer Skill	3rd	2nd	3rd	5th			III
Drivers License	Yes	Yes	Yes	No		II	
Knowledge Test Score	85	75	20	75	I	II	
Education	UG	UG	UG	UG	I		III

---

* Computer skill and bookkeeping skill are usually referred to as 1st, 2nd, . . . degree in Japan.

** Participants were informed that the Knowledge Test scores could range from 0 to 100.

*** "UG" denotes "university graduate."

Table 2

Mean Ratings of Desirability of Applicants as a Function of Condition, Member Type, and Phase of Measurement (Experiment 1)

		Pre-Discussion		Post-Discussion	
Condition	Applicant	I & II	III & IV	I & II	III & IV
C1	A	86.55	71.61	79.23	76.18
		(5.85)	(8.65)	(6.25)	(12.69)
	B	70.45	92.05	85.70	90.89
		(5.10)	(4.29)	(6.59)	(4.88)
	C1	35.93	65.27	34.43	49.45
		(10.33)	(9.02)	(16.70)	(19.52)
C2	A	85.09	72.25	79.64	77.77
		(9.12)	(8.28)	(8.76)	(13.92)
	B	64.80	89.68	75.09	80.86
		(7.86)	(4.31)	(11.23)	(11.60)
	C2	55.36	28.70	41.89	30.91
		(11.42)	(9.90)	(12.76)	(11.09)

Note. Numbers in parentheses are standard deviations. $N = 22$ for each cell.

Table 3

Frequencies of Group Decisions as a Function of Condition

Condition	Group Decision: Applicant A	Group Decision: Applicant B	Total
	Experiment 1		
C1	3	19	22
	(.14)	(.86)	
C2	13	9	22
	(.59)	(.41)	
	Experiment 2		
C1	2	13	15
	(.13)	(.87)	
C2	9	6	15
	(.60)	(.40)	

Note. Numbers in parentheses are the relative frequencies of the A vs. B choice for each condition.

Table 4

Mean Ratings of Desirability of Applicants as a Function of Condition, Member Type, and Phase of Measurement (Experiment 2)

		Pre-Discussion		Post-Discussion	
Condition	Applicant	I	II & III	I	II & III
C1	A	90.13	73.27	82.27	79.07
		(5.88)	(7.69)	(7.72)	(8.84)
	B	75.40	88.70	86.40	89.23
		(8.55)	(5.55)	(5.77)	(5.20)
	C1	33.47	55.90	38.33	43.63
		(12.99)	(10.10)	(18.39)	(16.48)
C2	A	88.87	73.53	85.13	83.30
		(6.15)	(4.70)	(5.76)	(5.06)
	B	75.00	86.33	79.13	83.57
		(7.16)	(5.42)	(8.02)	(5.06)
	C2	69.40	39.07	54.13	43.43
		(8.08)	(10.21)	(14.53)	(10.50)

Note. Numbers in parentheses are standard deviations. $N = 15$ for each cell.