The Means to an End:
Effects of Process on Group Choice and Problem Solving

by

Rebecca W. Hamilton

B.S., Policy Analysis
Cornell University, 1991

SUBMITTED TO THE SLOAN SCHOOL OF MANAGEMENT IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2000

© Rebecca W. Hamilton. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part.

Signature of Author:

Sloan School of Management
April 26, 2000

Certified by: Drazen Prelec
Professor of Management Science
Thesis Supervisor

Accepted by: Birger Wernerfelt
Professor of Management Science
Chairman, Department Committee
The Means to an End:
Effects of Process on Group Choice and Problem Solving

Rebecca W. Hamilton

Submitted to the Sloan School of Management
on April 26, 2000 in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

Abstract

The procedures used by groups to make choices and solve problems are critical: they define the opportunities group members have to evaluate others and influence their behavior; they determine when implicit knowledge and assumptions are shared. The three essays in this dissertation examine the role of intermediate outcomes - e.g., the alternatives being considered for a choice, the expected outcomes of an action with uncertain consequences, a diagram mapping relationships among concepts. While intermediate outcomes are often unobserved and implicit when individuals make choices or solve problems, in groups they are frequently observed and explicit. For example, a group might follow a decision process in which members generate a list of alternatives before making a final choice. Alternatives considered - even if not chosen - can affect both the final choice and group members’ satisfaction with the choice.

The first essay examines intuitive strategies people use to influence the outcomes of a group choice. Experimental evidence suggests people know how to propose alternatives that influence others’ choices, either by leveraging relationships among alternatives (e.g., making one alternative the “compromise” between two more extreme alternatives) or by eliminating alternatives others might prefer to one’s own favorite. Procedures are critical because they shape the context in which alternatives are evaluated.

The second essay suggests that people consider expected outcomes, such as having an equal chance to receive a preferred outcome, in addition to actual outcomes when assessing the fairness of choices with shared consequences. Experimental results show that factors such as the observability of the process change the relative importance of procedural fairness (expected outcomes) and distributional fairness (actual outcomes). One implication is that using appropriate procedures can improve joint outcomes.

The way people categorize concepts reflects the relationships they perceive among them. Thus, structured categorization techniques are often used to make tacit knowledge explicit. The last essay presents experimental results showing that individuals and groups categorize the same concepts in a qualitatively different manner. Moreover, the process of learning the technique in an individual or group setting can have lasting effects on how people perceive relationships among concepts.

Thesis Supervisor: Drazen Prelec
Title: Professor of Management Science
Acknowledgements

My warmest thanks go to the Marketing faculty at Sloan, for providing tremendous support and uncountable learning opportunities – including those at the lunch table – during my four years in the program. In particular, I would like to thank my advisors Drazen Prelec and Nader Tavassoli. Drazen gave me the freedom to choose a research topic I found fascinating and patiently guided its evolution, enthusiastically critiquing each new development and adding his sense of humor to the process. Nader was a extraordinary mentor, always taking the time to listen to my ideas, give me advice (usually right), or simply give me a hard time. I would also like to thank John Carroll for many helpful discussions, from my first semester as a Ph.D. student to my last. Finally, thanks to my fellow Ph.D. students in Marketing for both serious debates and enjoyable diversions.

Achieving balance in life is crucial, and I would like to thank my husband, my family and my friends both for supporting me in my work and for convincing me to spend some of my time away from it. Thank you for your love, patience, and perspective.
Contents

Abstract ........................................................................................................... 2
Acknowledgements .................................................................................. 3

Why Do People Suggest What They Don’t Want? Using Menus to Strategically
Influence Others’ Choices ........................................................................ 5
References .............................................................................................. 50
Appendices .............................................................................................. 52

When the Means Justify the Ends: Tradeoffs Between Distributional and
Procedural Fairness .................................................................................. 54
References .............................................................................................. 86
Appendices .............................................................................................. 89

Enduring Effects of Social Context on Categorizing Customer Needs .......... 93
References .............................................................................................. 129
Appendix .................................................................................................. 132
Why Do People Suggest What They Don’t Want?
Using Menus to Strategically Influence Others’ Choices*

* Thanks to John Carroll, Aimee Drolet, Rick Larrick, Drazen Prelec, Nader Tavassoli and seminar participants at MIT for discussions and helpful comments.
Why Do People Suggest What They Don’t Want?
Using Menus to Strategically Influence Others’ Choices

Abstract

Why would people suggest alternatives to others that they don’t want them to choose? Constructing comparisons for others can draw out tradeoffs among alternatives or create contrasts that influence others’ choices. For example, a salesperson might try to convince a customer to purchase a particular product by bringing the customer’s attention to a more expensive model, and arguing that the first product is a bargain in comparison. Because these contrasts and tradeoffs can systematically affect others’ choices, the person suggesting alternatives for a choice has a great deal of power to shape the decision process and the outcome.

Would people ever use such strategies to influence choices involving people they like and with whom they will continue to interact? Many choices made by groups begin with one person suggesting alternatives to others: for example, a family member might suggest several vacation ideas, or an organizational buyer might construct a list of potential suppliers for a product. This paper presents experimental evidence suggesting that people intuitively know how to construct menus that influence others’ choices, either by leveraging the relationships among alternatives or by eliminating competing alternatives from a menu. These strategies are demonstrated to be effective in influencing others’ choices, even when others are aware that the menus have been constructed strategically. In fact, for group choices with shared consequences, believing that menus had been created strategically by another group member actually enhanced rather than limited the effectiveness of the menus. Moreover, evidence from choices made by intact groups suggests that people use these strategies spontaneously, and that they are willing to use such strategies for consequential choices involving their friends.
Introduction

Anecdotal evidence suggests that strategically constructing a menu of options is a recognized technique for influencing others' choices. In his memoirs of his years spent advising Nixon, Henry Kissinger describes his attempts to revise the National Security Council’s official process for presenting options to the president. Because Cabinet departments were inclined to guide the president toward their preferred policy, yet were required by the National Security Council’s process to present three alternatives, the majority of the alternatives presented to the president were undesirable options. Kissinger (1979) explains that when “forced to present [multiple] options, the typical department will present two absurd alternatives as straw men bracketing its preferred option – which usually appears in the middle position.”

Strategic influence of this type has obvious implications for marketing, for example, in the construction of retail assortments, displays and product lines. Simonson and Tversky (1992) suggest that salespeople often exploit comparisons between products, and that a “common tactic used to convince consumers to purchase a given product is to present another product and argue that the former is a bargain in comparison with the latter.” Such tactics need not have purely negative implications: relationships among alternatives, such as price/quality tradeoffs, may serve to facilitate choice among the alternatives in the set even if the comparison alternatives are not seriously considered (Shocker, Ben-Akiva, Boccara & Nedungadi, 1991). For example, a realtor might make the purchase decision easier by showing a client an unaffordable home or one that is in disrepair (Shocker et al., 1991). Consider this communication with the Vice President of Sales at a graphic design supply firm:¹

“One of the products we sell is color copiers. ... When we train our salespeople, we tell them to always show the customer two or three copiers, not just one. This lets them show the customer the tradeoffs among the features of the products and compare price vs. features. ... There is one [copier] in particular that we emphasize, because it makes sense for most of our customers based on the features they get for the price. ... They show one of our lower end models with it and one of our higher end models with it.”

These real world examples suggest that people understand context effects intuitively. Might such influence tactics be used in everyday interactions among group members and negotiators? Negotiations often begin with one side making an initial offer, which can influence the other party’s perceptions of the possible outcomes (Bazerman & Neale, 1992). Similarly,

¹ Personal communication with Larry Salomon, VP of Sales for Charrette at Logan Airport, Sept. 2, 1999.
group decisions often begin with one group member suggesting alternatives to others: one person might recommend a few restaurants for a dinner with friends, a family member might suggest possible vacation spots, or an organizational buyer might construct a list of potential suppliers for a product. Techniques to influence others’ perceptions of desirable outcomes might be useful, for example, if one person wanted to convince some friends to go to a restaurant slightly more expensive than their usual meeting places. Would the person’s friends be more likely to choose the expensive restaurant when choosing between this restaurant and a few even more expensive alternatives (which make this restaurant look reasonable by comparison) than when comparing this restaurant to other establishments in their usual price range?

If people know how to create menus that strategically influence others’ choices, this phased process of narrowing down alternatives could give one group member great influence over the final choice made by a group. But if everyone understands these effects, why should they continue to work? One reason such a strategy might continue to work even if everyone anticipates its use is that the menu itself provides cues about the appropriate range of alternatives to consider (e.g., will it be a nice dinner or the usual?). Thus, just as the initial offer in a negotiation is assumed to convey information about the other party’s position, the menu of options a group member suggests might provide information about his or her preferences. Given that the choice has shared consequences – and their own digestion is likely to be enhanced by the satisfaction of their dining partners with the restaurant – others may be receptive to cues about the proposer’s preferences. This paper will investigate some of the conditions under which the knowledge that menus were created strategically enhances rather than limits the effectiveness of the menus in influencing others’ choices.

Three related questions are addressed in a series of experiments. First, do people know how to construct menus that will influence others’ choices? Second, if they have this knowledge, do people actually use such strategies to influence others’ choices when they are given the opportunity? Finally, are these strategies effective, and how do others’ beliefs about the manner in which a menu was constructed influence their evaluations of the alternatives and their choices from the menu? Three studies using hypothetical choices assess subjects’ knowledge of these effects, both in creating the menu and in choosing from the menu. In a fourth study, subjects made real choices with shared consequences for themselves and their friends to test their willingness to use these strategies in real group settings.
Strategic Influence Strategies

People encounter others’ attempts to influence them on a daily basis. Coworkers and family members, advertisers and sales people all send a variety of persuasive messages. Friestad and Wright (1994) suggest that such persuasive attempts are similar in that they involve two relevant roles, that of the persuasion agent and that of the persuasion target. Friestad & Wright (1995) suggest that through experience and learning, people develop a ‘folk knowledge’ of persuasion that they use to both interpret and construct persuasive messages. Because of this learning process, these authors argue that persuasion knowledge may be best evaluated as “knowledge in use,” or demonstrated effectiveness in convincing others (Friestad & Wright, 1999). Consistent with this framework, this paper will examine both persuasion knowledge and effectiveness by examining persuasion attempts and perceptions of persuasive attempts in real group settings.

Influence in group interactions has been conceptualized primarily in terms of verbal communication about the alternatives being considered (Meyers & Brashers, 1999). However, there are many forms of influence that shape decision outcomes. For example, the framing of a message provides a perspective for the listener, focusing attention on particular issues (Sussman, 1999). Either the range of alternatives considered for a choice (e.g., Simonson & Tversky, 1992) or the order in which they are considered (Levine & Plott, 1977) can influence decision outcomes by making different tradeoffs among alternatives more salient. These less traditional forms of persuasion, including framing, context effects and agendas, will be reviewed here.

Traditional forms of persuasion, in which individuals are given new information that affects their beliefs, have been distinguished from framing, which causes individuals to weight information they already have differently (Nelson, Oxley & Clawson, 1997). In research on political communications, a frame is the central idea that organizes the presentation of an issue or news story that defines the problem and identifies the relevant considerations. Similarly, a group member who suggests alternatives to others often has great power to shape the context in which the alternatives will be evaluated. For example, negotiators often choose the context of their messages (is a counteroffer presented as a gain from my original position or a loss from your offer?) as well as the characteristics of their message (the amount of the counteroffer; Blount & Larrick, 2000).
Previous research has identified systematic differences in outcomes based on the way a
game is framed (Larrick & Blount, 1997). Blount and Larrick (2000) study whether, when given
a choice among frames, people choose the frame that is likely to be the most profitable for them.
Using an interpersonal bargaining game, the authors find that their subjects are not very good at
choosing the frame that maximizes their own payoffs, due to misconceptions about the other
player’s responses, and perhaps also to concerns about fairness and norms. These results suggest
that both people’s knowledge of strategic influence techniques and their motivation to use such
techniques must be examined.

Context effects occur when the probability of choosing a particular alternative is
influenced by its relationships with other alternatives in the choice set (Simonson & Tversky,
1992), and these effects have been repeatedly demonstrated to influence individual choice. For
example, the compromise effect describes an increase in the choice probability of a middle
alternative, even if it is not superior to either of the more extreme alternatives that surround it
(Simonson, 1989). Verbal protocols from laboratory experiments testing context effects suggest
that subjects may have an intuitive understanding of such effects: subjects who made
compromise choices were quoted as saying “I’ll take the middle alternative” or “[I’ll take] the
one that is a combination of the two” (Simonson, 1989). While discussions of managerial
implications in papers on context effects often suggest ways that managers can use their
knowledge of these effects to develop retail assortments and product lines (e.g., Simonson &
Tversky, 1992), few researchers have examined the actual use of such tactics. This paper will
examine whether people leverage such context effects when presenting menus of options to
others for a group choice.

Another form of strategic influence relevant to group decision making is agenda setting.
If the decision process is modeled as a tree, an ‘agenda’ is a set of constraints on the order of
selecting or eliminating alternatives which determines the order in which the nodes are
considered (Hauser, 1986). For example, a top-down agenda selects first a category (e.g., type of
restaurant) and then a choice from within the category (e.g., a particular establishment), while a
bottom-up process compares features of each alternative across categories (comparing a favorite
restaurant of one type to a favorite restaurant of another type).

Blending personal and scientific motives, Plott and Levine (1978) propose a model of
agenda influence on group decision making, and describe their own experiences with agenda
influence on their flying club’s selection of planes (Levine & Plott, 1977). With some knowledge of the preferences of other group members, the authors were able to sequentially structure the decisions the club made to increase the probability that their own preferred outcome would be selected. This was achieved by forcing members to make early decisions which constrained their later choices of alternatives. Plott and Levine (1978) argue that an agenda affects the outcome in two ways: first, by limiting the information that each group member has about the dynamic patterns of preference within the group, and second, by determining which strategies are available to individuals. Their results underscore the importance of testing these strategies in real group settings, in which group members often have detailed knowledge about others’ preferences. Group members must not only have the knowledge of strategies to influence others, but also the knowledge of others’ preferences required to implement the strategies.

**Constructing Menus to Influence Choices**

This section presents hypotheses in three areas. The first two hypotheses concern knowledge: do people understand how to construct menus that will influence others’ choices from the menus? The third hypothesis addresses motivation: when given the opportunity, do people attempt to influence others’ choices by constructing menus strategically? Finally, the last hypothesis concerns effectiveness: do beliefs about how a menu was constructed influence others’ evaluations of the alternatives, and when are others most likely to choose as they believe they are expected to choose?

**Knowledge**

From the perspective of the person proposing alternatives for a choice, two strategies might be effective in influencing others’ choices. First, if the proposer assumes that others’ preferences for the alternatives are determined exogenously (rather than influenced by the decision context), the probability that others will choose a particular alternative from a menu should be increased by eliminating competing alternatives. The proposer might influence choice by constraining the menu so that others will choose substitutes for more preferred alternatives that are not available from the menu. Thus, examining the composition of the menu should reveal that from the proposer’s point of view, the alternatives competing with the favored alternative are less preferred than they would be if the proposer were not trying to influence the choice.
**Hypothesis 1:** A menu designed to influence others’ choices should compare the targeted alternative to other alternatives believed to be *less preferred* by others.

In contrast, a second potentially effective strategy leverages relationships among the alternatives proposed. Recognizing that relationships among the alternatives can influence choices from the menu, the proposer might structure the menu to favor the choice of a particular alternative. For example, the proposer might suggest both a more expensive restaurant and a less expensive restaurant to make a place he or she prefers appear to be a compromise between these two more extreme alternatives, increasing the probability that others will choose the preferred restaurant. In this case, the proposer expects to influence others’ *evaluations of the alternatives* as well as their choices.

Based on previous research, it is unclear whether people will be able to use the compromise effect to their advantage when creating menus for an interactive group choice. Given that people tend to choose compromise alternatives (Simonson, 1989; Simonson & Tversky, 1992), will the menu itself be composed of alternatives that were compromises within the larger set of available alternatives, or will proposers be able to take the perspective of others who will choose from the menus they construct? There is some evidence that for individual choices, the compromise effect applies to the process of narrowing down alternatives (forming a consideration set) as well as to making choices. For example, in an experimental investigation of the effects of new brand entry, Lehmann and Pan (1994) showed that brands that became extreme after new brands entered the market were *less* likely to be included in consideration sets, while brands that became compromises were *more* likely to be considered. However, subjects in their study had no incentive to imagine how others would choose from their consideration sets.

Whether the compromise effect can be used strategically to create menus for an interactive group choice depends on the proposer’s ability to imagine the menu as other group members will see it. If proposers recognize, for example, that others tend to avoid extreme alternatives, they might attempt to strategically influence the outcome of a group choice by positioning the targeted alternative as a compromise between two more extreme alternatives.

**Hypothesis 2:** A menu designed to influence others’ evaluations of the menu alternatives should leverage a specific *pattern of relationships* among the alternatives proposed.

---

2 While the compromise effect may be enhanced when the decision maker expects to justify responses to others (Simonson, 1989), compromise alternatives are not necessarily favored in group decisions. Simonson and Glazer (1995) show that managers avoid compromise alternatives for strategic choices whether working individually or in groups.
Hypotheses 1 and 2 reflect two different mechanisms for influencing choices from menus. Hypothesis 1 predicts that taking preferences as given, proposers will surround a targeted alternative with unattractive alternatives, increasing the likelihood that others will choose the targeted alternative instead of more preferred but unavailable alternatives. In contrast, Hypothesis 2 is based on the belief that relationships among alternatives may influence others' choices from the menu. The form of influence predicted by Hypothesis 1 can be distinguished from that predicted by Hypothesis 2 by examining the proposer's perceptions of others' preferences and the positioning of the preferred alternative within the menu. For example, acting in accordance with Hypothesis 2, a proposer might surround a target alternative with two more extreme (but not necessarily the most unattractive) alternatives. This menu would be consistent with Hypothesis 1 only if the two non-targeted alternatives were also the least attractive alternatives available.

Another way to distinguish between these two strategies is to examine the proposer's predictions about others' choices from a menu they have created. If the proposer expects to influence others' choices but not their evaluations of the alternatives, then the proposer should predict choice probabilities consistent with predictions based on his or her perceptions of others' preferences (e.g., using a logit model that assumes preferences for each alternative are independent). In contrast, if the proposer expects to influence others' evaluations of the alternatives based on their position within the menu, then predicted choice probabilities for the targeted alternatives should overstate calculations based only on others' perceived preferences.

By making such comparisons empirically, this paper will attempt to gauge the relative use of strategies in which the proposer attempts to limit others' choices or leverage relationships among alternatives.

Motivation

Does familiarity with strategic influence tactics mean that they will be used in real group interactions, or are there normative pressures in group situations that might preclude the use of such strategies? When given the opportunity, will proposers attempt to influence others' choices using the strategies outlined in Hypotheses 1 and 2?

One way to test whether strategic influence plays a role in the construction of menus is to vary the proposer's degree of control over the final choice from a menu. Suppose that a group is faced with a choice for which consequences are shared, such as a movie or restaurant. If the
proposer knows that someone else will make the final choice, and believes that he or she can effectively constrain the menu of choices available, the proposer might try to influence the outcome by comparing a preferred choice to less attractive alternatives (Hypothesis 1). In contrast, if the proposer expects to choose from the menu him or herself, it would be irrational to constrain the available choices early in the process by including unattractive alternatives in the menu (Kreps, 1988). Similarly, the proposer might try to influence others’ evaluations of the alternatives by giving a favored alternative the middle position in the set (Hypothesis 2). Constraining the final choice in this manner is irrational if the proposer either expects to choose from the menu him or herself, or if the proposer believes that others’ choices cannot be influenced.

Thus, when the proposer's degree of control over the final choice is either very high (e.g., if the proposer will choose from the menu him or herself) or very low (e.g., if the final choice will be made randomly from the menu), the proposer should be less likely to construct menus strategically, because strategic behavior will merely increase the probability of a poor choice. Holding other factors constant, such as the ability to gather information during the choice process, an intermediate degree of control should lead to the most strategic behavior. The proposer should use such strategies in constructing the menu when they have the greatest potential benefit and the lowest potential cost – for example, when the proposer believes that someone else will be choosing from the menu, but can accurately predict how this person will respond to the available alternatives.

**Hypothesis 3:** Menus will be constructed more strategically when the proposer expects someone else to choose from the menu than when the proposer expects to choose from the menu him or herself.

Systematic differences between menus created for oneself and others to choose from will suggest that proposers are in fact attempting to influence others’ choices based on the composition of the menu. If the composition of a menu is more strategic (in the manner predicted by Hypotheses 1 and 2) when the proposer has less control over the final choice, this suggests that proposers are balancing the costs (probability that a less preferred alternative will be chosen) against the benefits (probability that the most preferred alternative will be chosen) of constructing a menu strategically.

---

3 This distinction is likely to be a matter of degree. For example, if others will see a menu, the proposer might
Effectiveness

Will these strategies effectively influence others’ choices? If the proposer can actually constrain the menu from which others choose, then there is little doubt that substituting less preferred for more preferred alternatives to compete with a target alternative (Hypothesis 1) will increase the probability that the target will be chosen. Effectiveness simply requires accurate prediction of others’ preferences. For example, if only one group member suggests restaurants, and this person knows that others prefer Chinese restaurants to Indian restaurants, and Indian restaurants to Italian restaurants, he or she could effectively convince others to choose an Indian restaurant by limiting the options to Indian and Italian restaurants.

When the proposer has less control over the menu, however, context effects (Hypothesis 2) might be more effective, because the menu itself can shape others’ evaluations of the alternatives. For example, if other group members can suggest additional alternatives, the proposer can no longer eliminate competing alternatives from the menu, but can still enhance the desirability of a particular alternative by defining its local context. It has been demonstrated that choices often depend on the menu of alternatives presented to an individual (e.g., Simonson, 1989; Simonson & Tversky, 1992). This has been explained in various ways: by the rules an individual uses to decide among alternatives (e.g., the rule ‘never choose the largest slice of cake that you are offered;’ Sen, 1997), by the ease of justifying one’s choices (Simonson, 1989), or by the informational value of the menu itself (e.g., the menu might allow one to make inferences about the average preferences of a market segment; Wernerfelt, 1995; Prelec, Wernerfelt & Zettelmeyer, 1997). This section will examine the composition of menus in relation to each of these arguments.

Whether beliefs about how the menu was constructed (e.g., were the alternatives selected by someone else acting strategically, or were they selected randomly?) will alter the effectiveness of context effects is an empirical question. If context effects are explained by compliance with principles or rules, then knowing how the menu was constructed should not change the effect (to be polite, one will still choose the second largest slice of cake). However, if context effects are due to the informational value of the menu, knowing how the menu was constructed might undermine the effect. For example, if people infer different things about average preferences when they see a store’s inventory at the beginning of the season (most constrain the menu to increase others’ satisfaction with his or her own final choice.
popular items should be best stocked) and at the end of the season (most popular items should be lowest in availability), their purchases may be influenced less by the actual inventory than by their inferences about average preferences. Thus, context effects might not work in a group setting if knowing how a menu was constructed (e.g., knowing that the menu was created to strategically influence their choices) allows others to ‘unravel’ the menu and guard against context effects.

In contrast, if context effects work because people are influenced by the ease of justifying their choices, the ability to unravel the menu could make other group members even more susceptible to context effects. For example, if people choose compromise alternatives because they can easily justify their choices (Simonson, 1989), knowing that the proposer favors the middle alternative might provide yet another reason for choosing it. Moreover, for choices affecting a group, others might wish to consider the proposer’s preferences when making the choice. To the extent that the menu is believed to communicate the proposer’s preferences, others’ choices might be more influenced by the composition of the menu.

One way to distinguish between these mechanisms is to examine people’s evaluations of the same alternatives under varying beliefs about the menu’s construction. Those who believe that a menu was created to influence their choices should evaluate the same alternatives differently from those who have no reason to believe the menu was created strategically. For example, if others believe the proposer created the menu by including unattractive alternatives with the target alternative, they might give a more negative evaluation to alternatives they believe the proposer did not target than they otherwise might.

Hypothesis 4: Beliefs about the strategy used by the proposer to create a menu will cause people to evaluate the alternatives in a manner consistent with this strategy.

It is important to note that even if strategic influence is recognized, it may not decrease other group members’ satisfaction with either the choice process or the outcome. In fact, the menu may serve as a form of communication between the proposer and other group members. For example, the proposer might use the menu to communicate to others the relevant tradeoffs to be made (e.g., quality or features vs. price), and the relevant range of alternatives to consider. If others interpret the menu as a form of communication rather than as uninvited strongarming, then they might be more likely to choose as the proposer wants them to choose when they recognize that a menu has been created strategically.
To summarize, Hypothesis 4 predicts that the proposer’s perceived motives in constructing the menu will influence the chooser’s evaluation of the alternatives. If differences in the proposer’s perceived motives cause the chooser to evaluate the alternatives differently, this may influence choices from the menu.

Study 1

This study tests whether people intuitively understand how to create menus that influence others’ choices. To act strategically, a proposer must not only create menus in a systematic rather than random fashion, but also consider how others will respond to the menu he or she creates. If the proposer believes that others’ preferences for the alternatives are determined exogenously (e.g., based on the values of their attributes), then the proposer is likely to use a strategy of eliminating competing alternatives from the menu (Hypothesis 1) to increase the probability that a particular alternative will be chosen. In this case, assumptions about others’ preferences are crucial to interpreting the proposer’s behavior, because the alternatives others consider most unattractive will differ depending upon whether their preferences are normally distributed, uniformly distributed, or skewed. In contrast, if the proposer believes that others’ preferences can be influenced by the composition of the menu itself, then the menu should be constructed to leverage relationships among alternatives (Hypothesis 2), for example, comparing the targeted alternative with two alternatives that are more extreme in their attribute values.

Subjects

Eighty-four subjects were recruited at a public airport and randomly assigned to conditions. Subjects were not paid for their participation.

Procedure

Subjects were provided with several items in a product category and asked to identify a subset of items that would maximize the probability that others would choose a particular item from the subset. In one test, subjects were asked to select three out of five portable grills for a product display that would maximize sales of one of the grills. In a second test, subjects were asked to invite three of five suppliers to demonstrate their office copiers, with the goal of convincing their workgroup to choose a particular copier.

Portable Grills. In the first test, subjects were asked to imagine that they worked for an appliance store, and that their job was to set up product displays. After reading descriptions of five portable grills, subjects were asked to select three of the grills to create a display designed to
increase purchases of one grill (Grill 3), assuming the store made no profit on sales of the other grills. Product specifications were those originally used by Simonson and Tversky (1992) in a test of the compromise effect.\(^4\)

<table>
<thead>
<tr>
<th></th>
<th>Grill 1</th>
<th>Grill 2</th>
<th>Grill 3</th>
<th>Grill 4</th>
<th>Grill 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooking Area</td>
<td>160</td>
<td>220</td>
<td>280</td>
<td>340</td>
<td>400</td>
</tr>
<tr>
<td>Weight</td>
<td>4 lbs.</td>
<td>7 lbs.</td>
<td>10 lbs.</td>
<td>13 lbs.</td>
<td>16 lbs.</td>
</tr>
</tbody>
</table>

With these five grills, a total of six three-item menus including Grill 3 can be created. To increase the interpretability of the responses, the number of grills (and thus, menus) from which subjects could choose was limited. Fifty-five subjects were given a choice among four grills (either Grills 1, 2, 3 and 4 or Grills 2, 3, 4 and 5) and asked to choose three of the four grills, always including Grill 3, to display.

In the first condition, in which Grills 1, 2, 3 and 4 were available, subjects were told that Grill 5 was out of stock and were asked to choose two grills to display with Grill 3. First, suppose that others’ preferences are believed to be determined exogenously. To avoid losing market share to grills that are closer competitors, Hypothesis 1 predicts subjects will display Grills 1 and 4 with Grill 3 if they believe others’ preferences are normally distributed, but Grills 1 and 2 if they believe others’ preferences are uniformly distributed. In contrast, if others’ preferences are believed to be context dependent, Hypothesis 2 predicts that subjects will create displays that make Grill 3 the middle alternative in the set, selecting either Grills 2 and 4 or Grills 1 and 4.

In the second condition, in which Grills 2, 3, 4 and 5 were available, subjects were told that Grill 1 was out of stock and were again asked to choose two grills to display with Grill 3. In this condition, Hypothesis 1 predicts subjects will display either Grills 2 and 5 or Grills 4 and 5, while Hypothesis 2 predicts subjects will display either Grills 2 and 4 or Grills 2 and 5.

The remaining 29 subjects, assigned to the control condition, were told that both Grill 1 and Grill 5 were out of stock. To verify that there was not a systematic bias towards larger or smaller grills, they were asked to choose either Grill 2 or Grill 4 to display with Grill 3.

Office Copiers. A second test eliminated the constraint on available alternatives (all 5 alternatives were always available) and examined a series of two menus created by the subjects. Subjects were asked to imagine that their workgroup was in the process of purchasing a new

\(^4\) Simonson and Tversky (1992) compared choice probabilities for subjects who saw only one of three subsets of the
office copier. After reviewing five brands of copiers, they were asked to invite the suppliers of three of the brands to demonstrate their products with the goal of maximizing the probability that a particular target brand (Brand B or Brand D) would be chosen by their workgroup. Subjects were provided with the following data, based on product specifications from Xerox Corporation’s website:

<table>
<thead>
<tr>
<th></th>
<th>Brand A</th>
<th>Brand B</th>
<th>Brand C</th>
<th>Brand D</th>
<th>Brand E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Copies/Minute</td>
<td>12</td>
<td>14</td>
<td>18</td>
<td>24</td>
<td>33</td>
</tr>
<tr>
<td>Price</td>
<td>$1695</td>
<td>$1795</td>
<td>$2000</td>
<td>$2400</td>
<td>$3000</td>
</tr>
</tbody>
</table>

Each subject was asked to create two menus, one targeting Brand B and another targeting Brand D. This design provides the means to analyze whether subjects used consistent strategies to target two different alternatives, given that subjects should assume their workgroup’s underlying preferences are the same in both cases. Subjects were randomly assigned to one of the two conditions, which differed only by the order in which they created the menus (targeting Brand B first or targeting Brand D first). Subjects chose one of the six possible menus for each target brand.

*Figure 1: Market Share with Normal (1a), Uniform (1b), or Skewed Preferences (1c)*

![Diagram of market share](image)

What information does the selection of a given menu provide? Suppose that the target brand is B. Assuming that others’ preferences are not influenced by the menu, choosing either Brands A and E (if others’ preferences are perceived to be normally distributed, as in Figure 1a) or Brands D and E (if others’ preferences are perceived to be uniformly distributed, as in Figure 1b, or skewed, as in Figure 1c) would be consistent with Hypothesis 1. In contrast, choosing five grills, a set including grills 1-3, a set including grills 2-4, and a set including grills 3-5.
Brands A and C, Brands A and D, or Brands A and E, and predicting that others will be more likely to choose the middle alternative from the menu, would be consistent with Hypothesis 2. Because only the choice of Brands A and E is consistent with both Hypotheses 1 and 2, the crucial test for interpreting each subject’s strategy is to compare the sequence of two menus created.

To be consistent with Hypothesis 1, those who chose opposite extremes for the first menu (e.g., choosing D and E to target B, or choosing A and B to target D) must either choose opposite extremes for the second menu (if preferences are assumed to be uniformly distributed across alternatives) or extreme alternatives (A and E, if preferences are assumed to be skewed). Similarly, those who chose extreme alternatives for the first menu must choose extreme alternatives for the second menu (if preferences are assumed to be normally distributed), or opposite extremes (if preferences are assumed to be skewed) to behave consistently with Hypothesis 1. In contrast, the strategy of surrounding the target with non-extreme alternatives (e.g., choosing A and C to target B and then choosing C and E to target D) is consistent with Hypothesis 2, because subjects must believe that others are systematically more likely to choose middle alternatives.

**Results**

*Portable Grills.* It is crucial to the interpretation of the results to consider the assumptions subjects made about others’ preferences. The control condition verified that there was not a systematic bias in preferences favoring either smaller or larger grills. Fourteen of the 29 subjects (48%) chose to display Grill 2 to increase sales of Grill 3, and fifteen subjects (52%) chose to display Grill 4, suggesting non-skewed preferences.

In the two test conditions, subjects were asked to create a display of three grills to enhance sales of Grill 3. If subjects had chosen randomly among the three available menus, choices would have been divided evenly (33% share for each) among the three menus. However, a chi-square test shows that subjects’ menu choices differed significantly from random choices, $\chi^2(2) = 14.44, p < .001$. Moreover, subjects seemed to believe they could influence others’ evaluations of the alternatives by the way they constructed the menu. Thirty of the 55 subjects in the two test conditions (55%) chose to display Grills 2 and 4 to enhance purchases of Grill 3 (see Table 1), a choice which makes sense only if subjects believe others’ choices are influenced by
the context, as predicted by Hypothesis 2. In contrast, only 7 of the 55 subjects (13%) behaved in a manner consistent with only Hypothesis 1.

Table 1: Choice Sets Created to Increase Sales of Grill 3

<table>
<thead>
<tr>
<th>Condition</th>
<th>Consistent only with H1</th>
<th>Consistent with H1 or H2</th>
<th>Consistent only with H2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grills 1-4 Available</td>
<td>15% (1&amp;2)</td>
<td>44% (1&amp;4)</td>
<td>41% (2&amp;4)</td>
</tr>
<tr>
<td>Grills 2-5 Available</td>
<td>11% (4&amp;5)</td>
<td>21% (2&amp;5)</td>
<td>68% (2&amp;4)</td>
</tr>
<tr>
<td>Total</td>
<td>13%</td>
<td>32%</td>
<td>55%</td>
</tr>
</tbody>
</table>

Office Copiers. Again, subjects’ menu choices differed significantly from random choices, both when the targeted copier was Brand B, \( \chi^2(5) = 32.88, p < .001 \) and when the targeted copier was Brand D, \( \chi^2(5) = 47.95, p < .001 \). Almost all of the menu choices were divided among only three of the six possible strategies (see Table 2). Across the two choices, only 32 of 82 subjects (39%) used a strategy that could be interpreted as rational and consistent based on various assumptions about static preferences and market share (Hypothesis 1). Eleven subjects chose extreme alternatives (A and E) for both choices (rational if preferences are normally distributed), 9 subjects chose opposite extreme alternatives for both choices (choosing A and B when Brand D was the target, choosing D and E when Brand B was the target; rational if preferences uniformly distributed across alternatives), and 12 subjects chose extreme alternatives (A and E) for one choice and opposite extremes (A and B or D and E) for the other choice (rational if preferences are skewed).

Table 2: Menus Created to Increase Choices of Target Alternatives

<table>
<thead>
<tr>
<th></th>
<th>Consistent with only H1</th>
<th>Consistent with either H1 or H2</th>
<th>Consistent with only H2</th>
<th>Average of Other 3 Menus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target B</td>
<td>23% (D&amp;E)</td>
<td>29% (A&amp;E)</td>
<td>28% (A&amp;C)</td>
<td>7%</td>
</tr>
<tr>
<td>Target D</td>
<td>28% (A&amp;B)</td>
<td>30% (A&amp;E)</td>
<td>30% (C&amp;E)</td>
<td>4%</td>
</tr>
</tbody>
</table>

In contrast, 29 of 82 subjects (35%) used strategies that would be rational only if they perceived others’ preferences to be context dependent (Hypothesis 2).\(^5\) Eighteen subjects surrounded the target for both choices (A and C or A and D, then C and E or B and E), and 11 subjects chose extreme alternatives for one choice (A and E) and surrounded the target for the other choice.

---

\(^5\) Including the eleven subjects chose extreme alternatives (A and E) for both choices, 40 of 82 subjects (49%) chose menus that were consistent with the assumption that others’ preferences are context dependent.
Twenty-one subjects (26%) used strategies that did not support consistent assumptions about others’ preferences across the two menus. For example, 9 subjects chose opposite extremes for one choice (A and B when Brand D was the target and D and E when Brand B was the target) and surrounded the target for the other choice (C and E or A and C).

Discussion

It is difficult to make detailed comparisons between the results for grills and copiers because the set of available alternatives was constrained in the first test and because the positioning of the targeted alternatives within the original set differed across the two tests. However, both tests support a few general conclusions. Subjects’ menu choices for both portable grills and office copiers were decidedly nonrandom. Even when choices were unconstrained, subjects were most likely to construct menus that followed one of the two hypothesized strategies for influencing others’ choices. The results of Study 1 suggest that a large percentage of subjects are familiar with at least one normatively recognized strategy for influencing others’ choices from menus, and that individuals may vary in the strategies they use.

Strikingly, the choices of at least 55% of subjects who created menus for portable grills, and at least 35% of subjects who created menus for office copiers suggest that people believe they can influence others’ preferences by leveraging the relationships among alternatives in a set. The goal of Study 2 is to determine whether subjects correctly identified the menus from which others were most likely to choose the targeted alternatives.

Study 2

Creating a menu is only one step of the group choice process. To conclude that a menu can strategically influence the outcome of a group choice requires that the menu influences others’ choices in the desired manner. To test this, subjects in Study 2 were asked to make a choice from one of the menus that subjects in Study 1 could have selected. Subjects were either given information about three portable grills and asked to choose one of them, or they were given information about three office copiers and asked to choose one of them. The goal of this study was to evaluate the effectiveness of the menus created in Study 1 in influencing others’ choices.

Subjects

Seventy-four subjects were recruited at a public airport and asked to make a choice from a menu of three portable grills. Subsequently, one hundred thirty-four subjects were recruited at
a public airport and asked to make a choice from a menu of three office copiers. Subjects were not paid for their participation, and none of the subjects had previously participated in Study 1. 

Procedure

Portable Grills. Subjects were asked to imagine that they had decided to purchase a portable grill, and that they saw a display of three grills at the store they visited. The “display” each subject saw was one of the five possible menus (including Grill 3 and two other portable grills) that subjects in Study 1 could have created. Subjects were asked to identify which of the three grills they would be most likely to purchase. Subjects were randomly assigned to one of the five menu conditions. Fourteen subjects chose from the menu including grills 2, 3 and 4, and 15 subjects chose from each of the other four menus.

Office Copiers. Subjects were asked to imagine that the company they worked for was buying a new office copier to be shared by their workgroup. One of their colleagues had invited three suppliers to demonstrate their products, and the final choice from among the three copiers that had been demonstrated was now up to them. Each subject received one of the nine possible menus (including either Brand B or Brand D and two other copiers) that subjects in Study 1 could have created. Subjects were randomly assigned to one of the nine menu conditions. Fourteen subjects chose from the menu including grills A, D and E, and 15 subjects chose from each of the other eight menus.

Results

Portable Grills. Table 3 shows the percentage of subjects who chose each grill from each of the five menus of three portable grills. Across the five menus, 31% of the subjects chose the smallest grill in the set, 46% of the subjects chose the middle grill in the set, and 23% of the subjects chose the largest grill offered. In aggregate, these results are similar to those reported by Simonson and Tversky (1992) for a larger sample size.6

Subjects in Study 1 had been asked to create a menu that would convince others to choose Grill 3. This choice data provides a way to examine whether they were successful, and whether their menu selection was better than a random baseline. A quick glance at Table 3 shows that 85% of the subjects to whom Grills 1-4 were available selected menus from which Grill 3 was the grill most likely to be chosen, and 89% of subjects to whom Grills 2-5 were available

---

6 Simonson and Tversky (1992) report that when grills 1, 2, and 3 were available, 40% of 77 subjects chose grill 2, when grills 2, 3 and 4 were available, 44% of 70 subjects chose grill 3, and when grills 3, 4 and 5 were available, 47% of 72 subjects chose grill 4.
selected menus from which Grill 3 was most likely to be chosen. Based on the frequency with which each menu was proposed in Study 1, a weighted average of 50% of the subjects in Study 2 would have chosen Grill 3 (see Table 3).

Table 3: Choices of Portable Grills

<table>
<thead>
<tr>
<th>Grills Available</th>
<th>Grill 1</th>
<th>Grill 2</th>
<th>Grill 3</th>
<th>Grill 4</th>
<th>Grill 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grills 1, 2 and 3</td>
<td>40%</td>
<td>33%</td>
<td>27%</td>
<td></td>
<td>15%</td>
</tr>
<tr>
<td>Grills 1, 3 and 4</td>
<td>20%</td>
<td>60%</td>
<td>20%</td>
<td></td>
<td>44%</td>
</tr>
<tr>
<td>Grills 2, 3 and 4</td>
<td>29%</td>
<td>50%</td>
<td>21%</td>
<td></td>
<td>41%</td>
</tr>
<tr>
<td>Grills 2, 3 and 5</td>
<td>33%</td>
<td>53%</td>
<td>13%</td>
<td></td>
<td>21%</td>
</tr>
<tr>
<td>Grills 3, 4 and 5</td>
<td></td>
<td>33%</td>
<td>33%</td>
<td>33%</td>
<td>11%</td>
</tr>
<tr>
<td>Weighted Average (Unweighted Average)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>51% (46%)</td>
</tr>
</tbody>
</table>

In contrast, if the five menus had been selected with equal frequency, only 46% of the subjects in Study 2 would have chosen Grill 3. More formally, the menus that subjects in Study 1 actually selected resulted in a significantly higher probability that subjects in Study 2 would choose Grill 3 ($M = .50$) than if their selections had been randomly distributed across the five menus ($M = .46$; $t(107) = 2.21, p < .05$). This improvement over a random baseline suggests that subjects in Study 1 had good intuitions about how others would choose from their menus.

Office Copiers

Table 4 shows the percentage of subjects who chose each copier from the menu of three copiers they were given. Over all nine possible menus, 10% of the subjects chose the slowest copier in the set, 44% of the subjects chose the middle copier in the set, and 46% of the subjects chose the fastest copier offered, suggesting that in general, subjects preferred faster copiers.

As predicted, computing a weighted average based on the frequency with which subjects in Study 1 selected each of the six menus improves the success rate. The menus that subjects in Study 1 selected to target Brand B resulted in a marginally higher probability that subjects in Study 2 would choose Brand B ($M = .32$) than if their selections had been randomly distributed across the six menus ($M = .28$; $t(158) = 1.90, p < .10$). The same was true for Brand D. The menus that subjects in Study 1 selected to target Brand D resulted in a significantly higher probability that subjects in Study 2 would choose Brand D ($M = .44$) than if their selections had...
been randomly distributed across the six menus \((M = .41; t(158) = 2.20, p < .05)\). Again, this improvement in the results suggests that subjects in Study 1 had correct intuitions about how others would choose from their menus.

Table 4: Choices of Copiers

| Copier Brands Available | A  | B  | C  | D  | E  | Study 1 Menu Choices
|-------------------------|----|----|----|----|----|----------------------
|                         |    |    |    |    |    | Target B | Target D |
| A, B, C                | 7% | 40%| 53%|    |    | 28%       |
| A, B, E                | 0% | 40%|    | 60%|    | 29%       |
| B, C, E                | 13%| 27%|    | 60%|    | 3%        |
| A, B, D                | 7% | 47%| 47%|    |    | 10%       | 28%      |
| B, C, D                | 13%| 47%| 40%|    |    | 7%        | 4%       |
| B, D, E                | 14%|    | 53%| 33%|    | 23%       | 8%       |
| A, C, D                | 13%| 60%| 27%|    |    | 0%        |
| A, D, E                | 14%|    | 43%| 43%|    | 30%       |
| C, D, E                |    | 13%| 40%| 47%|    | 30%       |
| Weighted Average (Unweighted Average) | 32% (28%) | 44% (41%) |

Discussion

At this point, it is useful to summarize the results of the first two studies. Study 1 showed that people seem to intuitively understand two hypothesized techniques for constructing sets of alternatives that influence others' choices, and that they are able to apply these techniques when asked to do so. Study 2 demonstrated that the subjects in Study 1 consistently performed better than randomly in identifying which menus would influence others' choices in the desired manner. However, the results of Studies 1 and 2 provide little insight into the causal mechanisms for these effects. What were the assumptions subjects in Study 1 made about others' preferences, and what were the inferences subjects in Study 2 drew from the menus they received?

It is important to examine these causal mechanisms to determine when these strategies might be used in real group choices. In contrast to the unidentified others for whom the menus were constructed in Study 1, a group member constructing a menu might have quite detailed information about other group members' preferences. As in standard tests of context effects, subjects in Study 2 had no reason to believe the menus had been created to strategically influence
their choices. However, in the context of a group choice, those choosing from a menu might consider the motives of the proposer in constructing the menu. Study 3 examines both the creation of menus and choices from these menus in the context of a choice being made by a group, and collects detailed information about both the proposer’s and the chooser’s beliefs.

**Study 3**

Knowledge of others’ preferences and concern about both their preferences and inferences are important influences on choices made by and for a group. For example, Plott and Levine (1978) caution that their agenda-setting would have been less effective if they had not had detailed information about others’ preferences. Their ability to influence their flight club’s purchase of planes was based on their knowledge about the biases of certain members and coalitions within the club. Thus, it is crucial to study choices for which these considerations are relevant. One way to do this is to study choices made by ‘intact’ groups, made up of people who already know each other.

Studying intact groups permits examination of whether subjects’ perceptions about specific other people’s preferences are consistent with the strategic influence tactics they use. In Study 1, subjects could rely only on their knowledge of the general public’s preferences to construct their menus, because the people who would later choose from their menus were not identified. In contrast, in Study 3, subjects created menus for the other members of their own groups, and subjects rated others’ preferences for each of the available alternatives. Comparing the menus subjects created to their perceptions of others’ preferences provides a link between their assumptions about others’ preferences and menu composition.

Another important factor in studying strategic influence in group choices is knowing how the chooser’s beliefs about how menus were created influence their evaluation of the menu alternatives (Hypothesis 4). If one group member believes that another group member created a menu strategically, he or she might perceive non-targeted alternatives to be less preferred than if the menu had not been created strategically, consistent with Hypothesis 1. Alternatively, those choosing from a menu might believe that another group member wants them to select the middle alternative from the menu, consistent with Hypothesis 2. In this study, subjects were asked to play two different roles, first creating menus from which others would choose, and then choosing from menus they believed others had created for them.
This study replicates Studies 1 and 2 in a group setting, allowing measurement of the proposer’s perceptions of others’ preferences, the proposer’s predictions about others’ choices, and the chooser’s perceptions of the proposer’s preferences. To manipulate their beliefs about how menus were constructed by other group members, subjects were given three slightly different sets of instructions for creating their menus. Then, in the second phase, subjects chose from a menu that they believed another group member had created for them. In actuality, however, all subjects chose from one of two standard menus, providing the ability to examine whether the subject’s beliefs about how the menu had been constructed influenced their evaluations of the alternatives, their perceptions about the proposer’s preferences for the alternatives, and their choices from the menus.

Subjects

Ninety-six subjects formed 32 self-selected groups of three members before signing up to participate in the study. Subjects received movie tickets for their participation. After arriving at the session, the three members of each group were randomly assigned to conditions, with one member of each group in each condition.

Procedure

Experimental sessions for this study were divided into two phases, menu creation and choice.

Menu creation. In the first phase, subjects were asked to imagine that they had just moved into an apartment with the other two members of their group, and that they were shopping for several new appliances to furnish the apartment. Subjects were led to believe that they would each research two appliances before selecting a subset of the available alternatives to propose to the other two group members. In actuality, however, all three group members proposed alternatives for the same two appliances. Two pairs of appliances (television and air conditioner, microwave oven and stereo speakers) were counterbalanced so that subjects in half of the groups created menus for each pair of appliances. Subjects were given a set of five alternatives for each appliance, from which they were asked to create a menu of three alternatives.

7 The product specifications for the appliances used in this study (TV, stereo speakers, air conditioner, microwave oven) were based on stimuli used by Lehmann and Pan (1994) for their study of context effects. See Appendix A for the alternatives presented to subjects.
To manipulate subjects’ beliefs about how the menus they would choose from in the second phase, subjects were given three different sets of instructions in the first phase as they experienced the role of proposer. One third of the subjects were asked to create menus to maximize the probability that another member of their group would choose a specific alternative from the menu, as in Study 1 (Strategic Experimenter Goal condition). The second group of subjects was asked to create menus to maximize the probability that the other group members would choose the alternative they believed was the best choice for the group (Strategic Group Goal condition). Finally, as a control condition, one third of the subjects were asked to create a menu of the three alternatives they thought were the best for their group (Nonstrategic Group Goal condition). Subjects were not told that their own instructions differed from the instructions given to other group members.

After they had created their menus, subjects in all three conditions were asked to predict the probability that the other two group members would choose each of the three alternatives from the menus they had created. Next, subjects rated their own preferences for the five alternatives available as well as their perceptions of others’ preferences for the five alternatives.

Choice. In the second phase, subjects made a choice from a menu of three alternatives for two appliances they had not seen before. Subjects were led to believe they were choosing from menus that had been created by another member of their group in the first phase. In actuality, however, to increase the level of experimental control, exactly the same menus were distributed to all three members of the group (see Figure 2). It was important to remove all true variation among the menus because the goal was to test whether beliefs about how the menu had been created had affected preferences and choices. In the absence of true variation among the menus, comparisons between subjects across the three conditions (Strategic Experimenter Goal, Strategic Group Goal, Nonstrategic Group Goal) should reflect the degree to which preferences and choices were influenced by differing beliefs. Choices from menus and preferences for the three alternatives were compared across conditions to determine whether beliefs about how the menus were constructed had influenced choice and evaluation of alternatives.

Subjects were in fact misled twice, creating false beliefs with respect to two points. First, subjects were told that the menus they received had been created by another member of their group. Second, because subjects were not told that the other members of their group had been instructed differently during the first phase, they should have believed that the menu they
received had been created by another group member using the same instructions they had received. This assumption was verified by asking subjects to reconstruct the original set of five alternatives from which their menu of three alternatives had been created, to identify the alternative they believed the proposer expected them to choose from the menu, and to rate their own preferences for the five reconstructed alternatives for each appliance and their perceptions of the proposer’s preferences for the same five alternatives.

*Figure 2: Actual and Perceived Manipulation*

<table>
<thead>
<tr>
<th>PERCEIVED MANIPULATION</th>
<th>ACTUAL MANIPULATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Member 2:</td>
<td>Member 1:</td>
</tr>
<tr>
<td>• creates menu</td>
<td>• creates menu</td>
</tr>
<tr>
<td>• chooses from</td>
<td>• chooses from</td>
</tr>
<tr>
<td>Member 1’s menu</td>
<td>standard menu</td>
</tr>
<tr>
<td>Member 3:</td>
<td>Member 2:</td>
</tr>
<tr>
<td>• creates menu</td>
<td>• creates menu</td>
</tr>
<tr>
<td>• chooses from</td>
<td>• chooses from</td>
</tr>
<tr>
<td>Member 3’s menu</td>
<td>standard menu</td>
</tr>
<tr>
<td></td>
<td>Member 3:</td>
</tr>
<tr>
<td></td>
<td>• creates menu</td>
</tr>
<tr>
<td></td>
<td>• chooses from</td>
</tr>
<tr>
<td></td>
<td>standard menu</td>
</tr>
</tbody>
</table>

*Expected Results*

How should subjects respond to these instructions? Across the three conditions, two comparisons are particularly informative: a comparison between the Strategic Group Goal condition and the Strategic Experimenter Goal condition, and a comparison between the Strategic Group Goal condition and the Nonstrategic Group Goal condition.

Subjects should use the same influence strategies in the Strategic Group Goal condition and Strategic Experimenter Goal condition, though oriented toward targets selected by the subjects themselves in the Strategic Group Goal condition and toward preselected targets in the Strategic Experimenter Goal condition (see Table 5). Collecting both preference data and menu data allows a comparison of the degree to which menus are viewed primarily as a mechanism for constraining choices (e.g., by contrasting the target with more extreme and less preferred alternatives as in Hypothesis 1) or as a mechanism for influencing others’ preferences (e.g., by contrasting the target with high end and low end alternatives as in Hypothesis 2). Both the
proposer's perceptions of others' preferences and the composition of the menus will be examined to distinguish between the two influence strategies. Aside from the fact that the targeted alternatives should be less preferred in the Strategic Experimenter Goal condition because they are assigned rather than selected, the relationships between targeted and non-targeted alternatives should be the same in both conditions.

Table 5: Predicted Results Across Conditions

<table>
<thead>
<tr>
<th>Condition</th>
<th>Creation of Menu</th>
<th>Choice from Menu</th>
</tr>
</thead>
</table>
| 1. Strategic Experimenter Goal   | • Non-targeted alternatives less preferred and more extreme than condition 3 (H1).  
  |                                 | • More likely to position targeted alternative in the middle of the menu than in condition 3 (H2). | • Lower evaluation of alternatives not perceived to be targeted than in condition 3 (H4).  
  |                                 |                                                            | • Lower evaluation of alternative perceived to be targeted than in condition 2. |
| 2. Strategic Group Goal          | • Non-targeted alternatives less preferred and more extreme than condition 3 (H1).  
  |                                 | • More likely to position targeted alternative in the middle of the menu than in condition 3 (H2). | • Lower evaluation of alternatives not perceived to be targeted than in condition 3 (H4). |
| 3. Nonstrategic Group Goal       | • No difference between targeted alternatives or preferences for targeted alternatives in conditions 2 and 3. | • Menu alternatives perceived to be less extreme than in conditions 1 or 2.  
  |                                 |                                                            | • Less likely to identify middle alternative as the target than in conditions 1 or 2. |

The second important comparison is between the Strategic Group Goal condition and the Nonstrategic Group Goal condition. There should be no difference across the two conditions in the average ratings of the alternatives or in which alternative is believed to be the best choice for the group. However, the composition of the menus created should differ systematically across conditions. Consistent with Hypothesis 1, non-targeted alternatives should be less preferred in the strategic condition than in the nonstrategic condition. Consistent with Hypothesis 2, targeted alternatives should be more likely to take the middle position of the menu in the strategic condition than in the nonstrategic condition. Based on the strategic composition of the menu, subjects should predict higher choice probabilities for targeted alternatives relative to the other
two alternatives in the set in the Strategic Group Goal condition than in the Nonstrategic Group Goal condition.

In the second phase, the chooser did not see the original set of five alternatives. Due to differing beliefs about the manner in which the menu was constructed, Hypothesis 4 predicts that subject’s own preferences and perceptions of the proposer’s preferences for the three menu alternatives should vary across conditions. Essentially, the perceived composition of the menu in the second phase should reflect the composition of the menu the subject created in the first phase. For example, in the two strategic conditions, subjects should devalue alternatives not perceived to be targeted by the proposer.

When should subjects be most receptive to influence, and more likely to choose the alternative they think the proposer expects them to choose? How subjects will react to the notion of choosing from a menu that they believe has been created strategically is an empirical question that will be addressed in the next section.

Results

There were no differences across randomly assigned conditions in subjects’ expected frequency of future interaction with other group members (M = 8.20 on a scale 1-9, with 9 indicating the highest frequency), or past interaction (M = 7.40). There were also no differences across conditions or appliances in subjects’ perceived knowledge of others’ preferences.

Figure 3: Proposer’s Preferences for Alternatives Across Conditions

Menu creation. Preference ratings (on a scale of 1-9, where 9 is the most preferred) were extremely consistent for each condition across the four appliances. In all three conditions, the
proposer's preference ratings for the five available alternatives were lowest for the most extreme alternatives (highest and lowest priced appliances) and highest for the middle alternatives (see Figure 3).

Table 5 reports results combined across the four appliances for subjects’ ratings of their own preferences and their perceptions of others’ preferences. One unpredicted effect of condition was that subjects rated all five alternatives less favorably in the Strategic Experimenter Goal condition than in the two Group Goal conditions, perhaps reflecting dissatisfaction with the task. More importantly for our comparisons, however, there were no differences in average preferences for the alternatives across the two Group Goal conditions. As predicted, there were no consistent differences in which alternatives subjects identified as their most preferred (the target alternatives), nor were there differences in subjects’ preference ratings for these alternatives across the two Group Goal conditions (see Table 5).

Table 5: Preference Ratings

<table>
<thead>
<tr>
<th>Condition</th>
<th>Own Ratings</th>
<th>Others’ Perceived Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Target Alternative</td>
<td>Other Two Alternatives</td>
</tr>
<tr>
<td>Strategic Experimenter Goal</td>
<td>5.83(^{a,b})</td>
<td>4.13(^{b})</td>
</tr>
<tr>
<td>Strategic Group Goal</td>
<td>8.69(^{a})</td>
<td>4.48(^{c})</td>
</tr>
<tr>
<td>Nonstrategic Group Goal</td>
<td>8.52(^{b})</td>
<td>6.12(^{b,c})</td>
</tr>
</tbody>
</table>

- Ratings differed significantly between Strategic Experimenter Goal and Strategic Group Goal conditions (\(p < .05\)).
- Ratings differed significantly between Strategic Experimenter Goal and Nonstrategic Group Goal conditions (\(p < .05\)).
- Ratings differed significantly between Strategic Group Goal and Nonstrategic Group Goal conditions (\(p < .05\)).

As predicted by Hypothesis 1, proposers compared less preferred alternatives with the target in the Strategic Group Goal condition than in the Nonstrategic Group Goal condition. While average preferences did not differ across the two conditions, subjects’ perceptions of others’ preferences for the two non-targeted alternatives included in the menu were significantly lower in the Strategic Group Goal condition (\(M = 4.95\)) than in the Nonstrategic Group Goal condition (\(M = 6.03\), \(t(62) = 3.20, p < .01\)). This pattern suggests that subjects are comparing the

---

\(^{8}\) Ranks 1 through 5 were assigned to the available alternatives. For air conditioners, average rank of the target alternative was 3.38 in the Strategic Group Goal condition (SGG) and 3.50 in the Nonstrategic Group Goal condition (NGG), \(t(16) = .28, p > .78\). For microwaves, average rank was 2.25 in both conditions. For televisions, average rank was 2.56 in SGG and 3.38 in NGG, \(t(16) = 1.95, p = .06\). For speakers, average rank was 4.00 in SGG and 3.75 in NGG, \(t(16) = .81, p > .42\).
target with less preferred alternatives in the Strategic Group Goal condition to influence others’ choices from the menu (see Figure 4).

**Figure 4: Proposer’s Perceptions of Others’ Preferences for Menu Alternatives**

![Graph showing proposer's perceptions of others' preferences for menu alternatives.](image)

Within subject comparisons also provide support for Hypothesis 1 and show consistency across the two strategic conditions. While both the proposer’s own preferences and others’ perceived preferences for the non-targeted alternatives were significantly lower than average preferences in the Strategic Experimenter Goal condition (for own preferences paired \(t(31) = 7.24, p < .001\); for others’ preferences \(t(31) = 7.15, p < .001\)) and in the Strategic Group Goal condition (for own preferences paired \(t(31)=5.42, p<.001;\) for others’ preferences \(t(31) = 4.77, p < .001\)), the opposite was true in the Nonstrategic Group Goal condition. In fact, in the Nonstrategic Group Goal condition, the proposer’s own preferences and others’ perceived preferences for the non-targeted alternatives were significantly higher than average preferences (for own preferences paired \(t(31) = 2.23, p < .05;\) for others’ preferences \(t(31) = 2.52, p < .05\)).

Because the proposer prefers less extreme alternatives (see Figure 3) and believes that others prefer less extreme alternatives, Hypothesis 1 predicts that strategically created menus should include more extreme alternatives than nonstrategically created menus. Defining the lowest priced and highest priced alternatives (for product specifications, see Appendix A) as ‘extreme’ alternatives, the number of non-target alternatives that could be extreme for each proposal set was 0, 1 or 2. As predicted, for all four appliances, the two non-targeted alternatives were significantly more likely to be extreme alternatives in the two Strategic conditions than in
the Nonstrategic condition (see Table 6). Also as predicted, there was no difference in the number of extreme alternatives proposed in the two strategic conditions \((p > .80)\).

**Table 6: Number of Extreme Non-targeted Alternatives**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Television</th>
<th>Air Condition</th>
<th>Microwave</th>
<th>Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strategic Experimenter Goal</td>
<td>1.56(^a)</td>
<td>1.44(^a)</td>
<td>1.38(^a)</td>
<td>1.31(^a)</td>
</tr>
<tr>
<td>Strategic Group Goal</td>
<td>1.56(^b)</td>
<td>1.38(^b)</td>
<td>1.44(^b)</td>
<td>1.44(^b)</td>
</tr>
<tr>
<td>Nonstrategic Group Goal</td>
<td>.69</td>
<td>.81</td>
<td>.75</td>
<td>.81</td>
</tr>
</tbody>
</table>

\(^a\) Significantly higher in Strategic Experimenter Goal than in Nonstrategic Group Goal condition \((p < .02)\).

\(^b\) Significantly higher in Strategic Group Goal than in Nonstrategic Group Goal condition \((p < .01)\).

Consistent with Hypothesis 2, subjects were significantly more likely to create menus positioning the target alternative in the middle of the set in the two strategic conditions \((M = 66\%)\) than in the Nonstrategic Group Goal condition \((M = 47\%, t(94) = 2.74, p < .01)\). In contrast, there was no difference between the Strategic Experimenter Goal condition \((M = 70\%)\) and the Strategic Group Goal condition \((M = 61\%, p > .20)\) in the frequency with which the target alternative was the middle alternative.

Not only did subjects’ menus differ systematically across conditions, but subjects in the strategic conditions also believed that their menus would influence others’ choices. This is demonstrated by comparing the proposer’s predictions about others’ choices in the two group goal conditions. While there was no difference across conditions in which alternative was believed to be best for the group, subjects predicted a much higher probability that another group member would choose this alternative in the Strategic Group Goal condition than in the Nonstrategic Group Goal condition. In the Strategic Group Goal condition, subjects predicted an average choice probability of 57% for the most preferred alternative and 20% for each of the other two alternatives, while in the Nonstrategic Group Goal condition, they predicted an average choice probability of only 47% for the most preferred alternative and 27% for each of the other two alternatives (see Table 7).

Using a simple logit model\(^9\) that assumes static rather than context dependent preferences, subjects’ ratings of others’ preferences and the menus they created can be used to calculate choice probabilities for the targeted alternatives (see Table 7). In the Nonstrategic

\(^9\) Choice probabilities were calculated using the following formula: choice probability of \(x\) from set \(\{x,y,z\}\) = \(\exp(\text{others' pref for } x)/\left(\exp(\text{others' pref for } x) + \exp(\text{others' pref for } y) + \exp(\text{others' pref for } z)\right)\).
Group Goal condition, subjects’ predictions were no different from the calculated choice probabilities \((p > .83)\). In contrast, in the two strategic conditions, subjects’ predictions \((M = 57.48)\) significantly overstated the calculated choice probabilities \((M = 52.08, t(61) = 2.86, p < .01)\), suggesting that subjects believed others’ evaluations of the alternatives would be influenced by the menus they had created.

**Table 7: Predicted vs. Calculated Choice Probabilities**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Predicted Probabilities</th>
<th>Calculated Probabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Target Alternative</td>
<td>Other Alternatives</td>
</tr>
<tr>
<td>Strategic Experimenter Goal</td>
<td>55.43 (^a)</td>
<td>22.23 (^b)</td>
</tr>
<tr>
<td>Strategic Group Goal</td>
<td>59.04 (^a)</td>
<td>20.47 (^b)</td>
</tr>
<tr>
<td>Nonstrategic Group Goal</td>
<td>46.89</td>
<td>26.69</td>
</tr>
</tbody>
</table>

\(^a\) Frequencies significantly higher than in the Nonstrategic Group Goal condition \((p < .01)\).

\(^b\) Frequencies significantly lower than in the Nonstrategic Group Goal condition \((p < .01)\).

Replicating the results from Study 1, the data suggest that some people use a rule based on eliminating competing alternatives from the menu (consistent with Hypothesis 1) while others use a rule based on context effects (consistent with Hypothesis 2). Table 8 examines the consistency of subjects’ menus with each of these hypotheses.\(^{10}\) As predicted by Hypothesis 1, a significantly smaller proportion of menus created in the Nonstrategic Group Goal condition \((M = .16)\) eliminated competing alternatives than in the Strategic Experimenter Goal condition \((M = .48, t(62) = 4.23, p < .001)\) or the Strategic Group Goal condition \((M = .44, t(62) = 3.59, p < .001)\).

As predicted by Hypothesis 2, a smaller proportion of menus created in the Nonstrategic Group Goal condition \((M = .47)\) were consistent with the compromise effect than in the Strategic Experimenter Goal condition \((M = .70, t(62) = 2.92, p < .01)\) or the Strategic Group Goal condition \((M = .63, t(62) = 1.86, p < .10)\). Because some menus were consistent with both strategies, the third and fourth columns of Table 8 combine the two strategies, showing that subjects in the Nonstrategic Group Goal condition were significantly less likely to use either strategy than subjects in the two strategic conditions.

\(^{10}\) In the Strategic Experimenter Goal condition, if subjects had constructed their menus by randomly selecting two alternatives in addition to the targeted alternative, two of the six possible menus (33%) would be consistent with H1 and three of the six possible menus (50%) would be consistent with H2. Four of the six menus (67%) would be consistent with either H1 or H2.


Table 8: Strategies for Constructing Menus

<table>
<thead>
<tr>
<th>Condition</th>
<th>Consistent with H1</th>
<th>Consistent with H2</th>
<th>Consistent with H1 or H2</th>
<th>Not consistent with either H1 or H2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strategic Experimenter Goal</td>
<td>48%(^a)</td>
<td>70%(^a)</td>
<td>88%(^a)</td>
<td>12%(^b)</td>
</tr>
<tr>
<td>Strategic Group Goal</td>
<td>44%(^a)</td>
<td>63%</td>
<td>81%(^a)</td>
<td>19%(^b)</td>
</tr>
<tr>
<td>Nonstrategic Group Goal</td>
<td>16%</td>
<td>47%</td>
<td>53%</td>
<td>47%</td>
</tr>
</tbody>
</table>

\(^a\) Frequencies significantly higher than in the Nonstrategic Group Goal condition \((p < .05)\).

\(^b\) Frequencies significantly lower than in the Nonstrategic Group Goal condition \((p < .05)\).

**Choices.** How did subjects’ experiences in the first phase affect their beliefs about the menus they were given in the second phase? Recall that to eliminate any true variation in the menus, all subjects were given one of two standard menus of three alternatives. As a manipulation check, subjects were asked to reconstruct the original set of five alternatives based on the menu of three alternatives they were given. As shown in Figure 5, choosers rated their preferences for the five alternatives they believed had been available to the proposer. In all three conditions, their preferences for the reconstructed alternatives were very similar to the preference ratings proposers gave to the actual set of five alternatives (see Figure 3). Again, the most extreme alternatives are the least preferred.

*Figure 5: Chooser’s Preferences for Reconstructed Alternatives*

![Chooser's Preferences for Reconstructed Alternatives](image)

As expected, though they received the same menus and had similar beliefs about the set of five alternatives originally given to the proposer, subjects had different beliefs about the menus they had received, supporting Hypothesis 4. Table 9 lists the chooser’s preferences for...
their actual choice from the menu (Own Ratings, Actual Choice), the chooser’s preferences for the alternative believed to have been targeted by the proposer (Own Ratings, Predicted Choice), the chooser’s preferences for the two menu alternatives not believed to have been targeted (Own Ratings, Other Two Alternatives), and the chooser’s preferences for the five alternatives believed to have been available to the proposer (Own Ratings, Original Set). Table 9 also compares the chooser’s own preferences to their perceptions of the proposer’s preferences for the same alternatives (Proposer’s Perceived Ratings).

Table 9: Choices and Reconstruction of Original Sets

<table>
<thead>
<tr>
<th>Condition</th>
<th>Own Ratings</th>
<th></th>
<th>Proposer’s Perceived Ratings</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Actual Choice</td>
<td>Predicted Choice</td>
<td>Other Two Alternatives</td>
<td>Original Set</td>
</tr>
<tr>
<td>Strategic Group Goal</td>
<td>8.56(^a)</td>
<td>8.27(^a)</td>
<td>5.41</td>
<td>5.78</td>
</tr>
<tr>
<td>Nonstrategic Group Goal</td>
<td>8.22</td>
<td>7.66</td>
<td>5.78</td>
<td>5.84</td>
</tr>
</tbody>
</table>

\(^a\) Significantly higher in Strategic Group Goal condition (p < .05).
\(^b\) Significantly lower in the Strategic Group Goal condition (p < .05).

Though all subjects had received one of two standard menus, both subjects’ own ratings of the alternatives in the menu and their perceptions of the proposer’s preferences for the alternatives differed across conditions (see Table 9), as predicted by Hypothesis 4. The proposer’s perceived ratings of the two alternatives not believed to be targeted were significantly lower in the Strategic Group Goal condition (M = 5.61) than in the Nonstrategic Group Goal condition (M = 6.42, t(62) = 2.25, p < .05). This suggests that in the Strategic condition, subjects believed that proposers had included less preferred alternatives in the menus to influence their choices. These comparisons are especially compelling because the average ratings across all five alternatives perceived to be available to the proposer (the original set of five alternatives) did not differ across conditions (p > .45).

Also consistent with Hypothesis 4, those in the Strategic Experimenter Goal condition and Strategic Group Goal condition (M = .95) believed their menus included more extreme alternatives than those in the Nonstrategic Group Goal condition (M = .73, t(94) = 1.87, p < .10). Moreover, there was a significant correlation between the number of extreme alternatives subjects had used to create their menus in phase 1 and the number of extreme alternatives they
identified in the menus they received in phase 2 (Pearson \( r = .41, p < .001 \)). This correlation suggests that subjects believed others had used the same strategies they had used to create their menus in the first phase.

This result is important because two streams of research suggest that perceiving a greater distance between menu alternatives could influence choices from the menu. First, Simonson and Tversky (1992) suggest that people’s tendency to avoid extreme alternatives is one of the motivations underlying the compromise effect. Second, Cooke and Mellers (1998) showed that the spacing of attributes for previously experienced alternatives can modify the psychological value of changes to an attribute, influencing judgments of subsequently experienced alternatives. For these reasons, believing that a targeted alternative is being compared with more extreme alternatives is likely to affect preferences and choices.

*Figure 6: Chooser’s Preferences for Menu Alternatives*

In 3 out of 4 cases (\( M = 76\% \)), subjects chose as they thought the proposer wanted them to choose. When they believed that the proposer had created the menus strategically, in the Strategic Group Goal condition, subjects were actually marginally *more* likely to choose as they predicted the proposer wanted them to choose (\( M = 80\% \)) than in the Nonstrategic Group Goal condition (\( M = 67\%, t(62) = 1.81, p < .10 \)). Interestingly, subjects rated the alternative they believed the proposer had targeted *more favorably* in the Strategic Group Goal condition (\( M = \)

---

11 Subjects received a menu including alternatives 1, 2 and 4 for one appliance and a menu including alternatives 2, 4 and 5 for another appliance (see Appendix A), with the order of these two menus counterbalanced between subjects.
8.27) than in the Nonstrategic Group Goal condition \((M = 7.66, t(62) = 2.29, p < .05)\), as shown in Figure 6. Moreover, subjects were happier with their actual choices in the Strategic Group Goal condition than in the Nonstrategic Group Goal condition: subjects’ ratings of their actual choices were marginally higher in the Strategic Group Goal condition \((M = 8.56)\) than in the Nonstrategic Group Goal condition \((M = 8.22, t(62) = 1.88, p < .10)\).

As one would expect, there was a positive correlation between choosing as predicted and perceived agreement with the proposer’s preferences (Pearson \(r = .34, p < .001\)). Subjects whose preference ratings for all five alternatives correlated more closely with their perceptions of the proposer’s preferences (median split at Pearson \(r = .80\)) were significantly more likely to choose as predicted \((M = 84\%)\) than those who believed their own preferences were less correlated with the proposer’s \((M = 67\%, t(94) = 2.91, p < .01)\).\(^{12}\) However, after controlling for the correlation between the proposer’s and chooser’s preferences, the partial correlation between choosing as predicted and believing that the menu had been created strategically was still significant \((r = .23, p < .05)\). Subjects were more likely to choose as they believed others to expected them to choose when they believed that others had attempted to influence their choices.

The preceding analysis suggests that subjects’ perceptions of the alternatives were influenced by their beliefs about the strategies used by others to create the menus. Subjects were conscious that they were being influenced and yet they seemed willing to be influenced in most cases. Were subjects also influenced unconsciously, as has been demonstrated in previous tests of context effects (e.g., Simonson, 1989; Simonson & Tversky, 1992)? The design of Study 3 allows conscious and unconscious effects to be examined simultaneously, because two manipulations happened simultaneously. First, subjects’ beliefs about how the menu had been created were manipulated (conscious influence), and second, the menu subjects were given was manipulated (unconscious influence; as in traditional tests of the compromise effect, subjects were given either alternatives 1, 2 and 4 or alternatives 2, 4 and 5).

As discussed previously, the conscious manipulation of subjects’ beliefs significantly changed their preference ratings for the same alternatives, whether they were given alternatives 1, 2 and 4 or alternatives 2, 4 and 5. In contrast, the unconscious manipulation of the menus given to the subjects had the same effect for all three belief conditions; subjects in all three

\(^{12}\) Subjects who chose as predicted for both choices had an average perceived correlation of .78, which was significantly higher than those who did not choose as predicted for both choices, who had an average perceived correlation of .58 \((t(94) = 3.29, p < .001)\).
conditions were equally likely to be influenced by context effects. For example, of the three alternatives available to the chooser, the menu’s middle alternative was most often identified as the alternative the proposer wanted the subject to choose ($M = 53\%$, a proportion that is significantly higher than chance, $t(95) = 5.35, p < .001$), and this frequency did not differ across conditions ($p > .56$). The middle alternative was also disproportionately chosen ($M = 45\%, t(95) = 3.61, p < .001$), and there were no significant differences across conditions in the frequency with which subjects chose middle alternatives ($p > .43$).

**Table 10: Choices from Menus**

<table>
<thead>
<tr>
<th>Appliance</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>5</th>
<th>Cohen’s h</th>
</tr>
</thead>
<tbody>
<tr>
<td>TV</td>
<td>33%</td>
<td>46%</td>
<td>21%</td>
<td></td>
<td>h = .63, $p &lt; .01$</td>
</tr>
<tr>
<td>Microwave</td>
<td>21%</td>
<td>46%</td>
<td>33%</td>
<td></td>
<td>h = .63, $p &lt; .01$</td>
</tr>
<tr>
<td>Air Conditioner</td>
<td>29%</td>
<td>21%</td>
<td>50%</td>
<td></td>
<td>h = .35, $p &lt; .05$</td>
</tr>
<tr>
<td>Speakers</td>
<td>17%</td>
<td>54%</td>
<td>29%</td>
<td></td>
<td>h = .39, $p &lt; .05$</td>
</tr>
</tbody>
</table>

Thus, subjects were equally influenced by context effects in all three belief conditions, suggesting that subjects were not conscious of or could not prevent being influenced by context effects even when they believed the menus had been created strategically. Choices for all four appliances showed evidence of context effects (following the method of Simonson & Tversky, 1992). The results in Table 10 compare choices across the two standard menus given to subjects, with one menu including alternatives 1, 2 and 4, and a second menu including alternatives 2, 4 and 5. These results suggest that menus can influence preferences and choices on multiple levels, both consciously and unconsciously.

**Discussion**

Study 3 provided detailed evidence that people know how to influence others’ choices by strategically constructing menus, replicating Study 1 and supporting Hypotheses 1 and 2. Both the proposer’s perceptions of others’ preferences and the proposer’s predictions about others’ choices support the hypothesized mechanisms.

Moreover, based on the chooser’s preference ratings, their evaluations of the alternatives they received clearly depended on their beliefs about the motives of the proposer, supporting Hypothesis 4. There were significant differences in the chooser’s perceptions of the same menus
across conditions, and marginal differences in the chooser's stated preferences for the alternative chosen from the menu. Subjects were influenced both consciously (based on their beliefs about how menus had been created) and unconsciously (based on the actual composition of the menu).

The results suggest that believing that a menu was created strategically may enhance rather than limit the effectiveness of the menu in influencing choices. For the hypothetical choices in Study 3, people were quite willing to choose as they believed others wanted them to choose. In fact, when subjects believed that a menu had been created strategically with the goal of influencing them to choose the alternative believed to be best for the group, they evaluated this alternative more favorably than when they did not believe the menu had been created strategically. Even when subjects believed that another group member had targeted an alternative selected by the experimenter, the belief that he or she considered the other alternatives in the menu inferior may have convinced subjects to comply with what they perceived to be the proposer's expectations. Clearly, beliefs about the influence strategies used by others matter.

**Study 4**

The results of the first three studies suggest that people know how to construct menus that will influence others' choices, and that these strategies are effective in influencing others' choices, even when others believe that the menus were created strategically. However, it is not yet clear that people are willing to use such strategies to influence the outcomes of real group choices. The goal of Study 4 is to show that subjects might use menus to strategically influence the outcomes of choices affecting themselves and their friends when they have an opportunity (but have not been directed) to do so.

To provide insight into the relevance of these strategies for more naturalistic group choices, group members were asked to choose a movie to see with the other two members of their group. The procedure the groups used to select their movie gave each group member the opportunity to constrain – and perhaps influence – the movie choices of the other group members. Hypothesis 3 predicts that when the proposer expects others to choose from a menu he or she has created, the menu should be constructed more strategically, and should be more consistent with the predictions of Hypothesis 1 and Hypothesis 2 than a menu that the proposer expects to choose from him or herself.
Study 4 compares menus created by subjects anticipating that they will choose from the menu themselves with menus created by the same subjects for the same choice but anticipating that another group member will make the final choice from the menu. A lottery procedure was designed to give each of these two menus an equal probability of being used to make the group’s choice. This lottery procedure varied the subjects’ degree of control over the final choice yet controlled for the subject’s own preferences for the choice alternatives, the subject’s perceptions of others’ preferences, and the subject’s concern for others.

Subjects

All subjects who participated in this study had previously participated in Study 3. Ninety-six subjects formed 32 self-selected groups of three members before signing up to participate in the study. As payment for their participation in Study 3, subjects received tickets to see a movie with the other two members of their group, and they were asked to choose their movie using the procedures described in this section.

Procedure

Groups used a phased choice process to select their movies. First, subjects were given a list of twelve movies currently playing at local theatres (see Appendix B), and they were asked to indicate whether they had seen or heard of each movie and to rate both their own preferences and their perceptions of the other two group members’ preferences for each movie. Next, each group member created two menus that included three movies each.

To measure the effect of decision control on menu composition, subjects were asked to create two lists of three movies, a My Choice list and a Your Choice list. The order of creating these lists was counterbalanced between subjects. The My Choice list would be used if the subject was asked to make the final choice for the group from his or her own menu. The Your Choice list would be used if another group member was selected to be the decision maker for the group, but used the subject’s menu to make the final choice. Each of the three group members knew they had a 1 in 3 chance to be the decision maker for the group. As the decision maker, subjects had an equal chance of making the choice from their own list of three movies or from another group member’s list of three movies. Thus, Member 1 had a 33% chance to be the decision maker for the group, a 17% (.33*.50) chance of making the choice from his or her own My Choice list, an 8% (.33*.25) chance of making the choice from Member 2’s Your Choice list, and an 8% chance of making the choice from Member 3’s Your Choice list.
Regardless of which group member made the final choice and who provided the list, groups knew that they would have five minutes to discuss the three movies on the list before the decision maker chose from the menu. This discussion period controlled for both the observability of the list and the opportunity to gather information from others, both of which have been demonstrated to affect the alternatives proposed for a group choice (Hamilton, 1999). The only difference between the two conditions was in the level of decision control maintained by the proposer. Because the proposer maintained control over the final choice from the menu in the My Choice condition, it was less costly to propose alternatives preferred by other group members in the My Choice condition than in the Your Choice condition. Thus, while subjects should include their most preferred movie in both menus, Hypothesis 1 and Hypothesis 3 together predict that the other alternatives in the menu should be less desirable in the Your Choice condition than in the My Choice condition. To test this, subjects were asked to predict which movie they would choose from their My Choice set and which movie others would choose from their Your Choice set before they knew who would be the decision maker for the group or which menu would be used for the choice.

Results

There was no difference between subjects’ own preferences averaged over the list of twelve movies ($M = 5.38$) and their perceptions of others’ preferences averaged over the list of movies ($M = 5.40, p > .80$). However, both menus that subjects created reflected their own preferences more than others’ preferences. For the movies in the My Choice menu, subjects’ own preferences ($M = 7.24$) were significantly higher than others’ perceived preferences ($M = 6.43$, paired $t(95) = 6.82, p < .001$). Similarly, for the movies in the Your Choice menu, subjects’ own preferences ($M = 7.04$) were significantly higher than others’ perceived preferences ($M = 6.30$, paired $t(95) = 6.22, p < .001$).

Forty-seven percent of the subjects (45 of 96) proposed a different list of movies in anticipation of their own choice (My Choice condition) and in anticipation of another group member’s choice from the list (Your Choice condition). There was no difference in this effect across orders (Your Choice first, My Choice first; $p > .54$) or across Study 3 conditions ($p > .25$). Two lines of argument prevent this finding from being dismissed as a simple demand effect.

---

13 The fraction of subjects proposing different sets of alternatives across conditions in anticipation of a lottery between the sets is consistent with two previous experiments in which between 45% and 55% of subjects proposed different sets for three different choices in two different studies (Hamilton, 1999).
First, it is possible to distinguish between subjects who proposed different menus and subjects who proposed the same menus across conditions. For example, the average perceived correlation in group member preferences for the full set of twelve movies was significantly higher when the two menus were the same (Pearson $r = .59$) than when the menus were different (Pearson $r = .45$; $t(93) = 2.10, p < .05$). This suggests that people were more likely to try to influence others when they believed others’ preferences differed from their own. In contrast, there was no difference in the predicted frequency of future interaction ($p > .93$) or frequency of past interaction with other group members ($p > .42$) for subjects who proposed the same vs. different menus.

Second, differences between the menus were systematic rather than random. For example, subjects’ own preferences for the movies in the My Choice menu ($M = 7.24$) were significantly higher than their preferences for the movies in the Your Choice menu ($M = 7.04$, paired $t(95) = 2.51, p < .01$). Moreover, subjects perceived others’ preferences to be marginally higher for the movies in the My Choice menu ($M = 6.43$) than for the movies in the Your Choice menu ($M = 6.30$, paired $t(95) = 1.82, p < .10$). Because the only difference across conditions was the proposer’s degree of decision control, this difference suggests that subjects were adding less preferred alternatives to the Your Choice menu in an attempt to influence other group members’ choices from the menu. In contrast, if subjects believed others would simply choose randomly from the menus they created, including unattractive alternatives would only increase the probability of a poor choice for the group.

Adding confidence to this argument, the average of the subject’s own preferences for movies in the Your Choice menu was higher when the two menus were the same ($M = 7.27$) than when the menus were different ($M = 6.78$; paired $t(93) = 2.20, p < .05$), and the average of others’ perceived preferences for the movies in the Your Choice menu was higher when menus were the same ($M = 6.38$) than when menus were different ($M = 5.91$; paired $t(93) = 1.62, p = .11$).

Moreover, there is evidence that lower overall preferences for the movies in the Your Choice menu result from the elimination of competing (more preferred) alternatives, consistent with Hypothesis 1. Half of the subjects (18 groups) were asked to predict their own choices from the My Choice menu and the two other group members’ choices from the Your Choice menu. Fifty subjects (of 54 asked to do so) made these predictions. Ninety-eight percent (49 of 50) of those who predicted their own choices included the alternative they said they would choose for
the group in both the My Choice and Your Choice menus they created. Thus, differences in average preferences across menus were caused by changes in the other two alternatives being compared with the targeted alternative. This explanation is corroborated by comparing preferences for the target alternative and other two alternatives across the two conditions (see Figure 7).

*Figure 7: Subjects’ Preferences for Menu Alternatives*

How did the groups’ choices differ across conditions? Interestingly, the proposer ended up in a better position relative to other group members in the Your Choice condition than in the My Choice condition. While there was no difference between the proposer’s preferences \((M = 7.00)\) and others’ preferences for the movies actually chosen in the My Choice condition \((M = 6.95, p > .93)\), the proposer liked the movies chosen in the Your Choice condition \((M = 7.59)\) significantly more than others liked them \((M = 6.52, t(30) = 2.47, p < .01)\). This result suggests that the proposer was able to effectively influence others’ choices: though the proposer’s preferences for the movies in the Your Choice menu were on average lower, others chose the ‘right’ movies from the menus, and proposers were actually better off in the Your Choice condition than in the My Choice condition.

Van Dijk and Vermunt (2000) find a similar result when they compare outcomes between ultimatum games (in which the recipient has the power to refuse an allocation) and dictator games (in which the recipient has no power of refusal). Because allocations for dictator games were more favorable to the recipients, they concluded that for the recipients, it pays to be powerless. Analogously, in the present case, the proposer seemed to be more considerate of others’ preferences in the My Choice condition (when others had no power to influence the final choice) than in the Your Choice condition.
Discussion

The results of Study 4 suggest that the technique of strategically constructing menus of options explored in the first three studies is relevant to real group choices. Though they were not asked to influence others’ choices, subjects spontaneously modified the composition of the menus they created in response to changes in their level of control over the final choice from the menu. Moreover, these systematic differences were observed for choices for which strategic influence might be expected to be minimal. First, subjects were making real choices with shared consequences for themselves and their friends. Second, all of the group members were familiar with the unconstrained set of alternatives, because each of them had created their menus from the same list of twelve movies. Thus, influence attempts may have been minimized because they could be readily identified by others. Given that significant effects are observed even under these limiting conditions, the results of Study 4 suggest that the strategic influence strategies identified in the first three studies can materially influence the outcomes of real group choices.

Conclusion

The results of four studies demonstrate that strategic menu construction is an important consideration for groups making choices. First, the results suggest that people understand how to construct menus that will influence others’ choices from the menus. Second, people’s beliefs about how a menu was constructed influenced their evaluations of the alternatives included in the menu in a manner consistent with these influence strategies. Finally, the results suggest that people are willing to use these strategies for choices involving people they like and with whom they will continue to interact. When given the opportunity, almost half of those observed attempted to influence others’ choices by constructing menus strategically, even when the choice was consequential and their influence attempts were fairly obvious to others. Thus, there is support for all three parts of the strategic influence equation: knowledge, motivation and effectiveness.

The results from Studies 1 and 3 suggest that people understand context effects intuitively. Both studies demonstrated that people know how to create menus that will influence others’ choices, either by leveraging the relationships among alternatives or by eliminating competing alternatives from a menu. An interesting follow-up question is how these strategies might differ if the decision process were less structured, for example, if suggestions were made sequentially by various group members. Next steps include studying more dynamic situations, in
which multiple people propose alternatives, and situations in which the alternatives proposed might entice or dissuade group members from participating in an activity. It is also interesting to consider how the use of these strategies might be affected by observing others using them.

Moreover, Studies 2 and 3 demonstrated that these strategies effectively influenced others’ choices, even when others are aware that the menus have been constructed strategically. In fact, in Study 3, subjects were actually more likely to choose as they thought the proposer wanted them to choose when they believed that the proposer had created the menu strategically. These findings suggest that under some circumstances, people might regard menus as a source of information to guide their choices, providing cues about others’ preferences and about the appropriate range of alternatives to consider for a choice. Choices are likely to be easier when menus are created strategically (e.g., because the most similar competing alternatives are eliminated from consideration, or because it is simple to apply the rule of choosing the middle alternative). Thus, because these strategies can be cooperative, they need not rely on naivete among those choosing from the menu, allowing such strategies to be effective in the long run as well as in the short run.

Finally, in addition to evidence that people are familiar with these techniques, there is evidence that people are willing to use these strategic influence techniques for a consequential choice involving their friends. Almost half of the participants in Study 4 created different menus for their own choice and for another group member’s choice. An interesting question is whether the same subjects would behave strategically for a different choice or as part of another group. The studies in this paper spanned various contexts and types of groups (e.g., choice of an office copier for a workgroup, choice of appliances with apartmentmates, choice of movie with friends), showing the applicability of these influence strategies to a wide range of group choices, but future studies could examine differences in the use of strategic influence both across people and across various types of groups.

This experimental evidence is consistent with the example used to motivate the paper, Kissinger’s description of the process used by Cabinet departments to present options to the president. The attempts of the Cabinet departments to influence the president’s policy decisions by presenting “two absurd alternatives as straw men bracketing …[the] preferred option” demonstrate both knowledge of influence strategies and motivation to use them (Kissinger, 1979). While those constructing the menus might not have studied the strategies they used as
academically as Kissinger studied them, the systematic manner in which they constructed menus of options suggests that they recognized context effects at least unconsciously. As evidence for their motivation, the departments were willing to implement their knowledge to influence real policy decisions. Finally, there was at least some concern about the effectiveness of these strategies. If Kissinger and Nixon had not believed that these strategies were effectively influencing policy decisions, they would not have worked to revise the National Security Council’s official process for presenting options to the president.

Was Nixon conscious of being influenced by the menus of options presented to him? Based on Kissinger’s ability to identify and describe the tactics, Nixon was probably aware that the Cabinet departments were trying to strategically influence his decisions. However, the results of Study 3 suggest that Nixon could have been influenced both consciously and unconsciously. When presented with a subset of the possible alternatives, it is difficult to imagine which alternatives are missing. As in traditional tests of context effects, the salient dimensions are defined by the local context, even when decision makers are aware that menus have been created strategically. Thus, the more insidious influence may have resulted from the exclusion of attractive competing alternatives rather than the inclusion of unattractive and more extreme alternatives.

How could this process have been improved? One of the reasons this behavior occurred is that the National Security Council’s official process required that a single Cabinet department present three options to the president rather than a single recommended course of action. Two aspects of this process can be questioned. First, should multiple options be presented to the president rather than a single option, and second, should a single department be given jurisdiction over the policy area? Addressing the first point, even if the majority of options presented are undesirable options, proposing multiple alternatives provides more information than proposing a single alternative. Proposing multiple alternatives identifies the relevant tradeoffs to be made and communicates which alternatives are unattractive as well as which alternatives are attractive. Addressing the second point, rather than giving a single Cabinet department jurisdiction over an area, should the recommendations of several independent advisors have been compared? Because this structure would prevent the departments from using the strategy of eliminating competing alternatives from consideration, this change might have provided the president with more viable options to consider. However, the costs and benefits of
such a policy must be weighed. Efficiency, leveraging expertise, and minimizing conflict are also valuable goals for a group decision making process, all of which might be attained by giving a single department jurisdiction.

In conclusion, this research has demonstrated that strategically constructing a menu of options can be an effective technique for influencing others' choices. The most important contribution of this paper is to demonstrate that such menus are relevant not only to traditional marketing applications, such as the construction of retail assortments, product lines and displays, but also to decision making procedures used by groups, such as organizational buying centers, families, and friends making joint decisions.

This research adds to our knowledge of group processes and group decision making as well as to our knowledge of strategic influence strategies. Previous research on group decision making has focused on the last phase of choice, in which group members make a choice from among a defined set of alternatives. In contrast, this paper focuses on an earlier stage of the group choice process, in which group members identify alternatives to be considered for the choice. Because group choices are often initiated by one group member suggesting alternatives to others, it is important to understand the impact of strategically constructing menus. Conceptualizing this phase of the group decision making process as one in which a group member proposes a menu of options to others links research on group decision making to existing research on context effects, agendas, and framing.
References


Appendix A

Half of the subjects in Study 4 created menus for televisions and air conditioners. They were presented with the following stimuli:

<table>
<thead>
<tr>
<th>Picture Quality</th>
<th>TV 1</th>
<th>TV 2</th>
<th>TV 3</th>
<th>TV 4</th>
<th>TV 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lines</td>
<td>254</td>
<td>320</td>
<td>355</td>
<td>385</td>
<td>435</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Power</th>
<th>TVI</th>
<th>TV2</th>
<th>TV3</th>
<th>TV4</th>
<th>TV5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Watts</td>
<td>4500</td>
<td>5100</td>
<td>5400</td>
<td>5800</td>
<td>6500</td>
</tr>
<tr>
<td>Price</td>
<td>$250.00</td>
<td>$285.00</td>
<td>$300.00</td>
<td>$325.00</td>
<td>$350.00</td>
</tr>
</tbody>
</table>

The other half of the subjects in Study 4 created menus for microwaves and speakers. They were presented with the following stimuli:

<table>
<thead>
<tr>
<th>Power</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Watts</td>
<td>775 watts</td>
<td>880 watts</td>
<td>940 watts</td>
<td>1010 watts</td>
<td>1120 watts</td>
</tr>
<tr>
<td>Price</td>
<td>$99.00</td>
<td>$135.00</td>
<td>$155.00</td>
<td>$179.00</td>
<td>$219.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sound Quality</th>
<th>Set 1</th>
<th>Set 2</th>
<th>Set 3</th>
<th>Set 4</th>
<th>Set 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality</td>
<td>70</td>
<td>77</td>
<td>82</td>
<td>86</td>
<td>95</td>
</tr>
</tbody>
</table>

Product specifications were designed so that after being exposed to a similar pattern of stimuli in phase 1, subjects would not be able to infer the placement of the alternatives they were given in phase 2 based on the distance between alternatives. For example, a menu including alternatives 2, 4 and 5 cannot be distinguished from a menu including alternatives 1, 3 and 5 or a menu including alternatives 1, 2 and 4 based on the distance between alternatives.
### Appendix B

<table>
<thead>
<tr>
<th>Movie Title</th>
<th>Type of Film</th>
<th>Stars</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Beauty</td>
<td>Drama</td>
<td>Kevin Spacey, Annette Bening, Chris Cooper, Allison Janney, Thora Birch</td>
<td>R</td>
</tr>
<tr>
<td>Blue Streak</td>
<td>Comedy</td>
<td>Martin Lawrence, Luke Wilson, Dave Chappelle, William Forsythe, Peter Greene</td>
<td>PG-13</td>
</tr>
<tr>
<td>For Love Of The Game</td>
<td>Drama</td>
<td>Kevin Costner, Kelly Preston, John Reilly, Jena Malone</td>
<td>PG-13</td>
</tr>
<tr>
<td>In Too Deep</td>
<td>Action/Adventure</td>
<td>Omar Epps, LL Cool J, Stanley Tucci, Nia Long, Pam Grier, Hill Harper</td>
<td>R</td>
</tr>
<tr>
<td>Love Stinks</td>
<td>Comedy</td>
<td>French Stewart, Bridgette Wilson, Bill Bellamy, Tyra Banks, Tiffani-Amber Theissen</td>
<td>R</td>
</tr>
<tr>
<td>Mickey Blue Eyes</td>
<td>Comedy</td>
<td>Hugh Grant, James Caan, Jeanne Tripplehorn, James Fox, Joe Viterelli</td>
<td>PG-13</td>
</tr>
<tr>
<td>The Muse</td>
<td>Drama</td>
<td>Sharon Stone, Albert Brooks, Andie MacDowell, Jeff Bridges, Steven Wright</td>
<td>PG-13</td>
</tr>
<tr>
<td>The Sixth Sense</td>
<td>Thriller</td>
<td>Bruce Willis, Toni Collette, Haley Joel Osment</td>
<td>PG-13</td>
</tr>
<tr>
<td>Stigmata</td>
<td>Thriller</td>
<td>Gabriel Bryne, Patricia Arquette, Jonathan Pryce, Nia Long</td>
<td>R</td>
</tr>
<tr>
<td>Stir Of Echoes</td>
<td>Thriller</td>
<td>Kevin Bacon, Kathryn Erbe, Illeana Douglas, Liza Weil, Kevin Dunn, Conor O’Farrell</td>
<td>R</td>
</tr>
<tr>
<td>The 13th Warrior</td>
<td>Action/Adventure</td>
<td>Antonio Banderas, Omar Sharif, Diane Venora, Vladimir Kulich, Dennis Storholi</td>
<td>R</td>
</tr>
<tr>
<td>The Thomas Crown Affair</td>
<td>Action/Adventure</td>
<td>Pierce Brosnan, Rene Russo, Denis Leary</td>
<td>R</td>
</tr>
</tbody>
</table>
When the Means Justify the Ends:
Tradeoffs Between Distributional and Procedural Fairness*

* Thanks to John Carroll, Dina Mayzlin, Drazen Prelec, Nader Tavassoli and Robert Zeithammer for discussions, statistical advice and helpful comments.
When the Means Justify the Ends: 
Tradeoffs Between Distributional and Procedural Fairness

Abstract

What does it mean to make a fair choice for a group? Does a fair choice require that all group members receive equivalent outcomes, or is a choice also fair when each group member has an equal opportunity to receive a preferred outcome? Consider a situation in which two people are trying to decide who should use the last available ticket to a sold-out event. Distributional fairness can be achieved only if neither person takes the ticket, yet joint satisfaction might be higher if one of them attends the event. The question addressed in this paper is whether people will be more tolerant of this distributional fairness if the allocation is made using a demonstrably fair procedure, such as a coin flip or a lottery.

Previous research has demonstrated that people care about the process by which outcomes are derived in addition to the outcomes themselves. Such a multidimensional perception of fairness is potentially significant because in some cases, distributional fairness is difficult to achieve, and in others, higher joint payoffs can be obtained only by tolerating greater distributional unfairness. If people make compensatory tradeoffs between distributional and procedural fairness, using a fair process to allocate unfair outcomes might increase satisfaction with distributionally unfair outcomes. Earlier research has not examined whether people make such tradeoffs because procedural fairness has been operationalized primarily as a categorical independent variable manipulated by the experimenter (e.g., comparing a fair trial procedure to an unfair trial procedure). In contrast, this paper introduces a more continuous, quantitative measure of procedural fairness, and allows subjects to determine the fairness of both the outcomes and the procedures used to make a choice.

Two studies show that distributionally unfair outcomes may be perceived as fair when they are achieved through a demonstrably fair process. Moreover, the studies show that both distributional and procedural fairness significantly predict fairness judgments and preferences for allocations among group members. Given the opportunity to choose the level of procedural fairness for a choice, people intuitively apply a criterion of equal opportunity when making allocations between themselves and others. Conditions under which procedural fairness is given greater weight, such as when the process is observable by others, are explored and normative implications are suggested.
**Introduction**

Two friends had been waiting in line together for 45 minutes, hoping to buy tickets for a sporting event that would certainly sell out. When they finally reached the ticket counter, they received the unhappy news that there was only one ticket left. Which of them would buy the ticket?

“Well, I guess we could flip for it,” said one.
“All right, that’s fair,” said the other.

But was it fair? No matter who won the coin toss, the outcome certainly would not be fair: one of them would buy the coveted ticket and the other would get nothing, even though both had invested an equal amount of time in line. How is it, then, that a coin toss practically symbolizes fairness in our culture? As a procedure, the coin toss has been institutionalized to the point where it is a legal and accepted means of deciding ties in statewide elections (CNN, 1998) as well as the rule of last resort for selecting teams for the SuperBowl playoffs.

This paper presents empirical evidence suggesting that the coin toss procedure has certain characteristics which may contribute to its special status: it is an observable, transparent process in which it is clear that both sides have an equal opportunity to achieve unequal outcomes. More generally, this paper makes the argument that the criterion of equal opportunity embodied in the coin toss is not only a normative consideration for allocating resources, but also a descriptive one, predicting group members’ preferences for allocating resources among group members. Further, it is hypothesized that people are more tolerant of distributional unfairness – greater differences in outcomes across group members – when the allocation is made using demonstrably fair procedures. For example, when the choice process can be observed by others, the criterion of equal opportunity should be given more weight relative to differences in group members’ outcomes than when the choice process cannot be observed.

**Distributional Fairness and Procedural Fairness**

What does it mean to make a fair choice for a group? Consider a situation in which each group member has an equal claim on some available resources (i.e., they have invested equally). Clearly, dividing resources equally is considered fair when the resources can be split into equal shares, such as when dividing a luscious dessert among guests at a dinner party. However, the allocation problem becomes more difficult when the resources cannot be divided equally, as when a group of friends with dissimilar preferences chooses a movie to see together. Does a fair
choice require that all group members receive equivalent outcomes (everyone predicts that they will like the movie equally), or is a choice also fair when each group member has an equal opportunity to receive a preferred outcome (a chance to see the movie they think they’ll like the most)? The perceived fairness of procedures like the coin toss suggests that fairness judgments might be multidimensional, sometimes made on the basis of equal opportunity – the fairness of the process – rather than equal outcomes.

Consider a somewhat more consequential decision. Suppose that a group of friends has decided to rent a condo near a ski resort for the weekend. Two condos are available for the same price, at an equal distance from the ski resort, with enough beds for everyone. One of the condos provides equivalent accommodations for everyone. The other condo has nicer common areas, but some of the bedrooms are palatial, while others are positively shabby. Suppose further that the group members will divide the rental costs equally. Which condo is more desirable overall might depend on how much the group values distributional fairness relative to higher average outcomes for the members, and on whether the group feels they will be able to fairly allocate the unfair outcomes (the palatial vs. the shabby accommodations). More specifically, to predict the group’s choice, it is necessary to know the relative importance they place on average outcomes, distributional fairness and procedural fairness.

All else equal, people prefer higher payoffs for themselves to lower payoffs. However, research has also shown that people value distributional fairness, and may sacrifice their own payoffs to achieve higher equity in payoffs (e.g., Bolton, Katok & Zwick, 1999; Camerer & Thaler, 1995; Keller & Sarin, 1988). Because distributional fairness is based on the relationship between one’s own outcomes and others’ outcomes, one way to test how much people value distributional fairness is to ask them to divide resources between themselves and others.

The technique of collecting preference ratings or rankings for various allocations has been used to measure the relative weights of absolute and comparative payoffs in utility functions. For example, Messick and Sentis (1985) find that an additive utility function that positively weights both the person’s own payoff and the difference between their own and others’ payoffs performs well in predicting preferences for allocations. Grzelak (1982) reports that a utility function of three additive components (subject’s own gain, other person’s gain, and the difference between the two) predicted 80-90% of subjects’ rankings of allocations. Loewenstein, Thompson & Bazerman (1989) found that in evaluating various allocations, most
people preferred equality to either advantageous or disadvantageous inequality, but not surprisingly, preferred advantageous inequality to disadvantageous inequality.

To the extent that people positively value distributional fairness, as demonstrated by these findings, the group in the previous example might be willing to give up some joint utility (e.g., nicer common areas in the ski condo) so that everyone has more equivalent accommodations. However, because people value fair procedures in addition to fair outcomes, another factor to consider is the process by which the accommodations will be allocated. It has been shown, for example, that litigants’ judgments about the fairness of a court trial might depend on both the outcome of the trial and the manner in which the trial was conducted (Thibaut & Walker, 1975).

Particularly when parties will continue to interact over time, great importance is placed on procedures. Research in negotiation has long distinguished between the actions each party takes during the negotiation and the outcomes of the negotiation (Valley, Neale & Mannix, 1995). In some cases, the negotiation process (how an individual is treated) may have more impact on the relationship than the outcome (how much utility was gained; Greenhalgh & Chapman, 1995). In groups, this interest in procedures can be explained rationally because the procedures used by the group affect the likelihood that group members’ interests will be protected in the long run (Thibaut & Walker, 1975). Even beyond the rationale of self-interest, some researchers have argued that the procedures followed by other group members provide symbolic information about one’s position and status within a group (Lind & Tyler, 1988).

The process used to arrive at the outcomes can alter people’s preferences for various allocations. For example, if the person who divides the resources is selected based on winning a game of skill rather than being selected randomly, divisions tend to be more favorable to the allocator because the allocator perceives his or her role as more legitimate (Hoffman, McCabe, Shachat & Smith, 1994). Kahnemen, Knetsch and Thaler (1986) also find significant effects of framing. In responding to scenarios pitting firms against consumers, people believed that it was fair for a firm to pass on a cost increase, but also felt that it was fair for firms to retain much of the margin when costs decreased. These results suggest that the manner in which allocations are made – not just the allocations themselves – influence assessments of fairness.

In the previous example, if the group agreed upon a fair process for allocating the more luxurious accommodations among the group members – such as by drawing straws or drawing
numbers out of a hat – could this compensate in some way for the unequal outcomes? That is, might people’s utility functions include a term for procedural fairness in addition to the term for distributional fairness that has been identified in previous research? Are these terms additive, so that higher procedural fairness might compensate for lower distributional fairness?

Several studies suggest that people derive utility somewhat independently from outcomes and procedures: for example, Walker, LaTour, Lind and Thibaut (1974) manipulated the procedural fairness and the verdict of a trial independently and found only main effects for both variables. In two different contexts, student evaluations of teachers and citizen evaluations of political leaders, Tyler and Caine (1981) also found significant main effects for both procedural fairness and outcomes, and no evidence for an interaction effect. Unfortunately, however, this research was unable to capture tradeoffs between distributional and procedural fairness because the degree of procedural fairness that characterized the choices was a categorical independent variable determined by the experimenter (e.g., comparing a fair trial procedure to an unfair trial procedure), rather than a dependent variable chosen by the subjects (e.g., Latour, 1978; Tyler & Caine, 1985; Lind, Kulik, Ambrose & de Vera Park, 1993; van den Bos, Lind, Vermunt & Wilke, 1997).

What if people could determine both the level of procedural fairness and the level of distributional fairness for a given allocation decision? Would people be more likely use a lottery mechanism to distribute unequal payoffs than to distribute equal payoffs? If they could use such a lottery mechanism, would they be more likely to choose outcomes that are distributionally unfair but allow the group to achieve higher joint payoffs? To address these questions, this paper empirically examines the manner in which people make tradeoffs between procedural fairness and distributional fairness.

The next section discusses how to measure distributional and procedural fairness for choices that affect a group. The hypothesis that people intuitively consider procedural fairness when evaluating allocations among group members is tested in Study 1. Next, a model of the decision maker’s utility in which he or she makes tradeoffs between distributional and procedural fairness is presented. Study 2 examines some of the conditions under which procedural fairness is given greater weight by decision makers, such as when the process can be observed by others and when there is little ambiguity about others’ preferences. This study was designed so that subjects could determine both the procedural fairness and the distributional
fairness of real allocations they made to themselves and their friends. The paper concludes by discussing some normative implications for achieving higher joint payoffs in groups.

**Evaluating Procedural Fairness**

In contrast to the numerous well-defined measures that have been developed to capture distributional fairness, few measures have been developed to reflect procedural fairness. Evaluations of fair procedures might be made using theoretically important criteria such as consistency across people, consistency across time, bias suppression, accuracy, correctability, and representativeness (Barrett-Howard & Tyler, 1986; Leventhal, Karuza & Fry, 1980). In empirical research, procedural fairness is often assessed using participant ratings of the fairness of various procedures (e.g., Lind et al., 1993; Tyler & Caine, 1985), or by using face-valid manipulations of fairness such as varying trial procedures (LaTour, 1978) or the accuracy of procedures (van den Bos, Vermunt & Wilke, 1997).

In other research, more objective, quantitative measures have been developed to address normative questions about procedures. The equity literature has distinguished between the equity of the actual outcomes of a choice (ex post equity) and the equity of the process used to arrive at these outcomes (ex ante equity; Keeney & Winkler, 1985). For example, if two people flip a coin to decide which of them wins a dollar bill, this process is ex ante equitable, because both people have an equal chance to win the dollar. However, neither of the resulting distributions is ex post equitable because in both cases, one person receives a dollar more than the other. Ex post equity would require an outcome giving each person fifty cents.

**Figure 1: Comparison of Two Lotteries**

<table>
<thead>
<tr>
<th>Lottery 1</th>
<th>.5</th>
<th>(1,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.5</td>
<td>(0,0)</td>
</tr>
<tr>
<td>Lottery 2</td>
<td>.5</td>
<td>(1,0)</td>
</tr>
<tr>
<td></td>
<td>.5</td>
<td>(0,1)</td>
</tr>
</tbody>
</table>

As another example, consider the two lotteries in Figure 1. Suppose that for each lottery, the payoffs of Person A and Person B are listed as (A, B), and the number on the appropriate
branch of the decision tree (in this case .5, or 50%) represents the probability that the outcome will occur. The two outcomes of Lottery 1, (1,1) and (0,0), are both ex post equitable because A’s and B’s payoffs are the same in both cases. In contrast, neither of the two outcomes of Lottery 2 are ex post equitable because for either outcome, one person receives a payoff one unit greater than the other. However, both lotteries are ex ante equitable because the expected payoffs for A and B are equal (.5).

A strength of the equity literature is that it facilitates more continuous comparisons (rather than binary or categorical comparisons) of procedures across lotteries and other situations in which outcomes are uncertain. Thus, measures for ex ante equity and ex post equity can be incorporated into social utility functions (e.g., Keeney & Winkler, 1985). However, one drawback of the equity literature is that it is based on external considerations, such as even distribution of risks, rather than on individual preferences and individuals’ perceptions of their standing relative to others (Boiney, 1995).

To address this concern, and to incorporate individual preferences for absolute and comparative payoffs, as has been done in the fairness literature, Boiney (1995) developed measures of ex post fairness and ex ante fairness based on the concept of envy (see Appendix A for a more detailed description of Boiney’s measures). In this context, envy is defined as the degree to which an individual would prefer another person’s payoff to his or her own payoff, and is computed as the average of the positive differences between others’ outcomes and one’s own outcomes. For example, for an outcome giving $5 to Person 1, $5 to Person 2, and nothing to Person 3, Person 1 and Person 2 have envy of 0, but Person 3’s envy is 5. According to Boiney’s measures, the ex post unfairness of this outcome is 16.67, which is the sum of the squared differences between each person’s envy and the average envy across people (in this case, 1.67).

Ex ante unfairness is computed as the sum of squared differences between each person’s expected envy and the average expected envy for all three people. For example, suppose that three people face the following lottery over outcomes, and that each of the three outcomes has an equal chance of being selected:

Outcome 1: $5 to Person 1, $5 to Person 2, and 0 to Person 3
Outcome 2: $5 to Person 1, 0 to Person 2, and $5 to Person 3
Outcome 3: 0 to Person 1, $5 to Person 2, and $5 to Person 3
In this case, even though the ex post unfairness of each outcome is 16.67 (as computed above), ex ante unfairness is 0. This is because the average expected envy across the three people, 1.67, is the same as the expected envy for each person. This example shows that ex ante unfairness can be 0 even when ex post unfairness is high. In fact, when there is no possibility of reallocation, ex ante unfairness is always less than or equal to the ex post unfairness of the possible outcomes.

While these measures provide a means for comparing the distributional and procedural fairness of lotteries from the perspectives of those who are affected by the outcomes, they have been conceptualized exclusively as normative comparisons. For example, Boiney's goal is to guide an external decision maker to the most fair and efficient choice. Thus, it is not known whether these measures capture procedural fairness in the sense that it is used in the procedural justice literature or in the negotiation literature.

In contrast, the concern of this paper is on describing people's behavior when allocating resources among themselves and other group members. The criterion of equal opportunity predicts that a measure of ex ante fairness like the one developed by Boiney will reflect people's assessments of fairness and preferences for outcomes, much as categorical manipulations of procedural justice have previously been found to predict people's assessments of fairness and preferences (e.g., LaTour, 1978; Tyler & Caine, 1981; Walker et al., 1974).

Hypothesis 1: People consider the ex ante fairness of allocations when evaluating the fairness of allocations and their preferences for allocations.

Previous research suggests that all else constant, equal outcomes are preferred to unequal outcomes (Loewenstein et al., 1989), and that all else constant, a fair process is preferred to an unfair process (Lind & Tyler, 1988). The contribution of this paper is to suggest that ex ante fairness can be used to translate concerns about procedural fairness into a form describing preferences over choices with uncertain outcomes. If this translation is possible, it would allow measurement of the tradeoffs people make between their own and others' payoffs, distributional fairness and procedural fairness. As a first step, to test Hypothesis 1, Study 1 examined whether

---

1 Boiney (1995) models fairness in decisions affecting a group, where the outcomes of various actions are uncertain and decisions are made by a third party. Measures of ex ante unfairness and ex post unfairness are developed as part of a broader index—a normative decision aid—to be used by an outside decision maker. Modified versions of these measures are presented here.
the ex ante fairness construct predicts people's assessments of fairness and preferences for lotteries.

**Study 1**

This study tested whether people consider the process by which an allocation is made when evaluating the fairness of the allocation and their preferences for the allocation. Subjects were asked to rate both the fairness of and their preferences for eight allocations which differed along four dimensions: their own expected payoffs, the expected payoffs of other group members, the distributional fairness of the allocation, and the procedural fairness of the allocation. The relationships between the subjects' ratings of the allocations and these four dimensions were analyzed.

**Subjects**

Fifty-nine undergraduate and graduate students participated in the experiment. Twenty-nine subjects completed one set of eight comparisons and thirty subjects completed a different set of eight comparisons.

**Procedure**

Subjects worked individually, but were asked to imagine that they were making choices for a group including themselves and two friends. Subjects were given the following instructions:

Suppose that you and two friends will be given some scratch-off lottery tickets. The lottery tickets come in denominations of $5, $3, $2 and $1.

You are the one who will decide how the tickets you receive will be distributed among the three of you. Please review the following pairs and indicate which of the two distributions, A or B, you think is more fair, and which of the distributions you prefer.

**Figure 2: Subjects Compared Allocations A and B**

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>You</td>
<td>Friend 1</td>
</tr>
<tr>
<td>100% chance</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>33% chance</td>
<td>5</td>
</tr>
<tr>
<td>Which do you think is more fair?</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>Which do you prefer?</td>
<td>A</td>
<td>B</td>
</tr>
</tbody>
</table>

63
For example, one comparison presented to the subjects is shown in Figure 2. Each allocation (A or B) was a list of outcomes, all of which had an equal chance of being the final outcome if that allocation was chosen. As in Figure 2, the probabilities were shown for each possible outcome. Each outcome was associated with a payoff to each of the three group members.

Subjects were asked to make eight such comparisons. As a group, subjects made fifteen different comparisons, which are listed in Table 1 using the format (subject’s own payoff, Friend 1’s payoff, Friend 2’s payoff). These fifteen comparisons were divided into two sets of eight comparisons, with one comparison being made by all subjects. Subjects were randomly assigned to one of the two sets of eight comparisons.

Results

The results, reported in Table 1, can be analyzed both at the level of the individual items and in aggregate.

Table 1: Results of Comparing Allocation A with Allocation B

<table>
<thead>
<tr>
<th>Allocation A</th>
<th>Allocation B</th>
<th>% Prefer A</th>
<th>% A More Fair</th>
</tr>
</thead>
<tbody>
<tr>
<td>(5,2,2)</td>
<td>(5,0,5),(5,5,0),(0,5,5)</td>
<td>86%</td>
<td>11%</td>
</tr>
<tr>
<td>(5,0,5)</td>
<td>(5,2,2)</td>
<td>31%</td>
<td>10%</td>
</tr>
<tr>
<td>(5,2,2),(2,5,2),(2,2,5)</td>
<td>(5,0,5),(5,5,0),(0,5,5)</td>
<td>61%</td>
<td>75%</td>
</tr>
<tr>
<td>(5,0,5),(5,5,0),(0,5,5)</td>
<td>(5,5,0)</td>
<td>30%</td>
<td>92%</td>
</tr>
<tr>
<td>(5,2,2)</td>
<td>(5,2,2),(2,5,2),(2,2,5)</td>
<td>79%</td>
<td>5%</td>
</tr>
<tr>
<td>(5,3,1)</td>
<td>(5,2,2)</td>
<td>29%</td>
<td>14%</td>
</tr>
<tr>
<td>(5,2,2),(2,5,2),(2,2,5)</td>
<td>(5,3,1),(1,5,3),(3,1,5)</td>
<td>70%</td>
<td>62%</td>
</tr>
<tr>
<td>(5,0,5),(5,5,0)</td>
<td>(5,2,2)</td>
<td>40%</td>
<td>45%</td>
</tr>
<tr>
<td>(2,2,5)</td>
<td>(5,0,5),(5,5,0),(0,5,5)</td>
<td>13%</td>
<td>18%</td>
</tr>
<tr>
<td>(5,2,2),(2,5,2),(2,2,5)</td>
<td>(5,5,0)</td>
<td>46%</td>
<td>90%</td>
</tr>
<tr>
<td>(5,0,5),(5,5,0),(0,5,5)</td>
<td>(5,0,5),(0,5,5)</td>
<td>100%</td>
<td>95%</td>
</tr>
<tr>
<td>(2,2,5)</td>
<td>(5,2,2),(2,5,2)</td>
<td>8%</td>
<td>0%</td>
</tr>
<tr>
<td>(1,3,5),(3,1,5)</td>
<td>(2,2,5)</td>
<td>41%</td>
<td>46%</td>
</tr>
<tr>
<td>(5,3,1)</td>
<td>(5,3,1),(1,5,3),(3,1,5)</td>
<td>63%</td>
<td>18%</td>
</tr>
<tr>
<td>(5,0,5)</td>
<td>(0,5,5)</td>
<td>96%</td>
<td>67%</td>
</tr>
</tbody>
</table>

Note: Subjects were told that each of the outcomes listed for an allocation had an equal chance of being the final outcome, e.g., 33% chance of (5,0,5), 33% chance of (5,5,0) and 33% chance of (0,5,5).

Clearly, subjects distinguished between the fairness of and their preferences for the allocations. For example, for the comparison shown in Figure 2, 86% of the subjects felt that

---

2 Because results for the shared item did not differ significantly across conditions, the results were aggregated for this item in Table 1.
allocation B was more fair, but 85% of the subjects preferred allocation A. Across the 15 comparisons, the allocation that was considered most fair was the same as the allocation that was the most preferred only 51% of the time (for 231 of the 450 items for which both questions were answered).

**Individual items.** Consistent with previous research, individual item comparisons showed that both subjects’ own payoffs and the distributional fairness of the allocations predict their fairness ratings and preferences for the allocations (e.g., Grzelak, 1982; Loewenstein et al., 1989; Messick & Sentis, 1985). In addition, these comparisons suggested that the procedural fairness of the allocations predicts fairness ratings and preferences.

Controlling for distributional fairness and procedural fairness, subjects preferred allocations that gave higher expected payoffs to themselves. For example, when comparing the allocations A: (5,0,5) and B: (0,5,5), 96% of the subjects preferred allocation A, which gave the subject a higher payoff than allocation B. However, when the subject’s own payoffs were held constant, the distributional fairness of the allocation was a deciding factor. For example, when comparing the allocations A: (5,3,1) and B: (5,2,2), 86% of the subjects rated allocation B as more fair, and 71% of the subjects preferred allocation B. The same pattern held for a different expected payoff to the subject and a different level of procedural fairness. When comparing the allocations A: {{5,2,2}, (2,5,2), (2,2,5)} and B: {{5,3,1}, (1,5,3), (3,1,5)}, 62% of the subjects rated allocation A as more fair, and 70% of the subjects preferred allocation A.

The individual item comparisons also make it clear that subjects are using the procedural fairness of the allocations to make their fairness judgments. When comparing the allocations A: {(5,5,0), (5,0,5), (0,5,5)} and B: (5,5,0), 92% of the subjects rated allocation A as more fair, though the distributional fairness of both allocations is the same. Similarly, when comparing the allocations A: (5,2,2) and B: {{5,2,2}, (2,5,2), (2,2,5)}, 95% of the subjects rated allocation B as more fair.

**Aggregate results.** Another way to analyze the results is to examine how the choices of the subjects in rating one of the allocations as more fair or more preferred varied with the differences in the subject’s expected payoff across the allocations, others’ expected payoffs across the allocations, the distributional fairness of the allocations, and the procedural fairness of the allocations. The ex post unfairness measure introduced in the last section was used to
operationalize distributional unfairness, and the ex ante unfairness measure introduced in the last section was used to operationalize procedural unfairness.

After computing the differences in the subject's expected payoffs (i.e., expected payoff for allocation A – expected payoff for allocation B), others' expected payoffs, ex post unfairness and ex ante unfairness for each comparison, a maximum likelihood probit estimation was run. Because each subject rated eight different comparisons, standard errors were adjusted for clustering within subjects (robust standard errors). Results using OLS, variance-weighted least squares regression, and standard probit (without adjusting the errors for clustering within subjects) gave very similar results. The results are reported in Table 2. Marginal effects (slopes) for each variable were computed by averaging the marginal effects across individual observations (Greene, 1997).

<table>
<thead>
<tr>
<th>Table 2: Aggregate Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preference for Allocation</td>
</tr>
<tr>
<td>$\chi^2(4) = 62.85, p &lt; .001$</td>
</tr>
<tr>
<td>Pseudo $R^2$</td>
</tr>
<tr>
<td><strong>Variable</strong></td>
</tr>
<tr>
<td>Own Expected Payoffs</td>
</tr>
<tr>
<td>Others' Expected Payoffs</td>
</tr>
<tr>
<td>Ex Post Unfairness</td>
</tr>
<tr>
<td>Ex Ante Unfairness</td>
</tr>
<tr>
<td>Intercept</td>
</tr>
</tbody>
</table>

These results show that subjects preferred allocations giving them higher expected payoffs and that they marginally preferred allocations giving others higher expected payoffs. With all other factors held constant at their mean values, subjects were 48% more likely to prefer allocations giving them their maximum expected payoff than allocations giving them their mean expected payoff, and they were 27% more likely to prefer allocations giving others their maximum expected payoffs than giving others their mean expected payoffs. Subjects also preferred allocations that were less ex post unfair and allocations that were less ex ante unfair. With all other factors held constant at their mean values, subjects were 19% more likely to prefer allocations with the mean level of ex post unfairness to allocations with the highest level of ex...
post unfairness, and they were 23% more likely to prefer allocations with the mean level of ex ante unfairness to allocations with the highest level of ex ante unfairness.

A similar pattern of results holds for fairness ratings. Subjects rated allocations as more fair when they expected higher payoffs for themselves, when others received higher expected payoffs, and when the allocations were less ex post unfair and less ex ante unfair. With all other factors held constant at their mean values, subjects were 66% more likely to rate an allocation giving them their maximum expected payoff as more fair than an allocation giving them their mean expected payoff, and they were 65% more likely to rate an allocation giving others their maximum expected payoff as more fair than an allocation giving others their mean expected payoff. They were 28% more likely to rate an allocation with the mean level of ex post fairness as more fair than an allocation with the highest level of ex post unfairness, and they were 33% more likely to rate an allocation with the mean level of ex ante fairness as more fair than an allocation with the highest level of ex ante unfairness.

**Discussion**

In support of Hypothesis 1, these results suggest that ex ante (procedural) fairness significantly predicts both subjects’ assessments of the fairness of allocations and their preferences for the allocations. The follow-up question, then, is whether people make tradeoffs between distributional fairness and procedural fairness. If a person’s utility function includes both ex post fairness and ex ante fairness as additive terms, as was assumed for this analysis, this is a question about the relative importance of each term in predicting preferences.

Under what conditions might the relative importance that subjects place on procedural justice and distributional justice be shifted? One variable that should shift the relative weights is the observability of the process. It has been demonstrated that when people don’t know how to evaluate the fairness of the outcomes, they often use the process as a heuristic to judge fairness (van den Bos, *et al.*, 1997). Empirical evidence also suggests that the opposite is true: people focus on distributive justice when they are unable to judge procedural fairness (*Leventhal, et al.*, 1980). Thus, procedural fairness might be given relatively more weight and distributional fairness relatively less weight when the process will be observed by others. In contrast, procedural fairness might be given relatively less weight and distributional fairness more weight when others can observe only the outcomes.
Observability

The process by which a choice is made may be particularly important for decisions made in the presence of others or on behalf of others (Sen, 1997). Procedural norms refer to normative expectations about appropriate processes to follow and appropriate treatment of others rather than prescriptions about outcomes (Lind & Tyler, 1988). For example, someone sensitive to procedural norms might be happy to take the most comfortable chair in the room or the largest slice of cake after winning it by lottery, but they would not choose it when offered a selection of chairs or slices. Observability, or the ability to monitor others’ actions, is an important consideration in group choice because the attributions that other group members make when an unfair outcome occurs depend on their understanding of why the outcome occurred. Different attributions are made about a person’s motives for taking the largest slice of cake when it is the only slice left on the plate than when it is the first to be taken.

Several authors have distinguished between the decision maker’s accountability for the outcome and accountability for the process (e.g., Doney & Armstrong, 1996; Beach & Mitchell, 1978). Process accountability is a consideration when the decision maker expects others to evaluate the quality of their decision making procedures (e.g., did enough effort go into the judgment process?). Being held accountable for the process may affect the process itself: for example, when decision makers expect others to evaluate the quality of their decision making procedures, they may over-gather information due to social pressure (Feldman & March, 1981). Similarly, knowing that others will observe the process might change the process that one uses to allocate resources among group members.

Several experimental manipulations have been used in an attempt to distinguish between true fairness and the appearance of fairness in making allocations (e.g., Hoffman et al., 1994; van Dijk & Vermunt, 1999; also see Camerer & Thaler, 1995). For example, an anonymity manipulation is often used to distinguish between behavior that is motivated by personal principles and behavior that is socially motivated (Reis, 1981). Hoffman et al. (1994) show that when making allocations, people behave less generously toward others when they know the experimenters will be unable to tell whether they behaved generously. If the motivation to use a fair allocation process is internal rather than external (e.g., if people follow a fair process for ethical reasons, or because they derive satisfaction from being fair even if no one else observes...
the process), then the social context in which the process is applied should not influence the process.

In contrast, if procedural fairness is norm-driven and others’ perceptions are important, then whether the process is public (observable by others) or private (unobservable) might affect the process. For example, Mikula and Schwinger (1978) suggest that people use allocation decisions as a “politeness ritual” which gives them the opportunity to demonstrate modesty, politeness and unselfishness. Because a public choice process allows others to evaluate expected outcomes in addition to actual outcomes, a procedural norm such as giving each person an equal opportunity to receive a preferred outcome can be enforced. Thus, if procedural fairness is extrinsically motivated, it should be given more weight relative to other considerations when the process can be observed by others.

_Hypothesis 2: The importance of procedural fairness will be greater when the process can be observed by others._

Hypothesis 2 suggests that procedural fairness will be given more weight when the process will be observed by others, but what are the implications for distributional fairness? The attributions that others make for outcomes can affect their tolerance for unfairness. For example, Blount (1995) demonstrated that preferences for fair outcomes are affected by beliefs about how the outcomes transpired (e.g., whether unfair outcomes were selected by a human or by a computer). She showed empirically that people were more willing to accept unfair outcomes generated by a computer (perceived to be unintentional) than those proposed by the other player (perceived to be intentional). Thus, attributing outcomes to environmental causes, such as random selection, rather than human causes, such as another person’s choice, might make people more tolerant of distributional unfairness. Perhaps this is why randomization is often used when equity is a concern but outcomes are unequally divisible, such as in draft lotteries (Keeney & Winkler, 1985).

If people believe that an allocation resulted from a fair process, such as a coin toss or lottery, more distributional unfairness might be tolerated. This is important because higher payoffs for the group as a whole might be achieved by allowing greater inequality among group members. For example, when the group of friends discussed earlier was choosing between two condos, the condo with unequal accommodations had nicer common areas. It was suggested that the group would be more likely to choose this condo if they could agree on a procedure to
allocate the accommodations fairly. In more general terms, people who know they can
demonstrate to others that they followed a fair process in allocating resources might be more
willing to consider alternatives resulting in distributionally unfair outcomes. Thus, groups might
be able to achieve higher efficiency because they are less constrained by attempting to achieve
distributional fairness.

**Hypothesis 3:** People will be more willing to consider alternatives resulting in
distributionally unfair outcomes when they can demonstrate to others that they used a fair
choice process.

Demonstrating to others that one has followed a fair choice process requires that others
be able to evaluate the fairness of the process. If procedural fairness is operationalized using the
ex ante fairness construct, based on differences in expected outcomes across group members,
then evaluating procedural fairness requires others to estimate the outcomes of the group
members for each alternative. Others are better able to evaluate these expected outcomes when
preferences for the alternatives are clear than when preferences are ambiguous. For example,
preferences for money are clear: it is known that everyone would prefer more money to less
money. In contrast, preferences for movies are more ambiguous: the distributional implications
of choosing an action/adventure movie for the group relative to a drama are harder to evaluate.

When preferences are clear, choice behavior is more interpretable and compliance with
procedural norms can be more easily monitored by others. Thus, if others can better interpret the
procedures being followed when preferences for the potential outcomes are clear, there should be
an interaction between the observability of the process and the ambiguity of preferences. That is,
observability should have a greater impact on procedural fairness when others’ preferences are
clear rather than ambiguous.

**Hypothesis 4:** Observability will have a greater impact on procedural fairness when
procedural fairness can be clearly evaluated by others.

**Tradeoffs Between Distributional and Procedural Fairness**

If people are more tolerant of unfair outcomes when they can see that the process is fair,
as the wide acceptance of the coin toss procedure suggests, will people who are proposing
allocations between themselves and others take this tradeoff into consideration?

The results of Study 1 suggest that the subject’s own payoffs, the payoffs of others, the
distribution of payoffs among group members (ex post unfairness) and the fairness of the
allocation process (ex ante unfairness) all significantly predict preference ratings for allocations. These results can be modeled as a simple additive utility function with four terms: (1) subject’s payoff: the expected value of the subject’s payoffs for the lottery, (2) other group members’ payoffs: the expected value of the average of other group members’ payoffs, (3) distributional fairness: $I_p(x)$, the average ex post unfairness for the alternatives in the lottery (see Appendix A), and (4) procedural fairness: $I_A(a)$, the ex ante fairness of the lottery (see Appendix A). The subject’s utility for lottery $a$, from which each outcome $x$ is selected with probability $p(x)$ and provides utility $u_i(x)$ to the subject and $u_j(x)$ to $(n-1)$ other group members can be written as follows:

$$\text{Utility of } a = \alpha \sum_{x \in a} p(x) u_i(x) + \beta \sum_{x \in a} \sum_j p(x) u_j(x)/(n-1) + \gamma \sum_{x \in a} p(x) I_p(x) + \delta I_A(a)$$

This utility function provides a structure to investigate whether the ability to demonstrate a fair decision process motivates subjects to focus more on procedural fairness and less on distributional fairness when the process is observable than when it is unobservable (Hypothesis 2 and Hypothesis 3). For example, suppose that the same subject proposes Lottery $a$ when the process is observable and Lottery $b$ when the process is unobservable. Based on revealed preferences, this difference indicates that Lottery $a$ gives higher utility than Lottery $b$ when the process is observable, and that Lottery $b$ gives higher utility than Lottery $a$ when the process is unobservable.

$$U_{\text{Public}(a)} - U_{\text{Public}(b)} > 0 \text{ and } U_{\text{Private}(b)} - U_{\text{Private}(a)} > 0$$

Such a change in preferences across conditions (observable process vs. unobservable process) suggests that the relative weights of these four terms differ across the two conditions. Previous research demonstrating that people are more generous in their allocations under less anonymous conditions (Hoffman et al., 1994), suggests that the relative weight on one’s own payoffs relative to others’ payoffs should decrease when the choice process is public. In addition, Hypotheses 2 and 3 predict that when allocations are observable, the relative weight on distributional fairness relative to procedural fairness should decrease. Study 2 examined these predictions by analyzing changes in the lotteries created when the choice process can be observed by others and when the choice process cannot be observed by others.

---

3 If allocations are not zero sum, this term captures a preference for higher joint payoffs.
Study 2

This experiment investigated whether the process of making a choice for a group is influenced by the observability of the choice process. The experimental procedures allowed subjects to determine both the level of distributional fairness and the level of procedural fairness of their choices. Subjects in Study 2 created lotteries similar to those evaluated by subjects in Study 1. They created lotteries that had real consequences for themselves and two friends, allocating instant lottery tickets among the three group members and choosing a single movie that their group would see together.

This study isolated the effect of observability by comparing the composition of lotteries that subjects expected others to observe with lotteries for the same choice that subjects did not expect others to observe. A public/private manipulation is often used to test concerns about others’ impressions of one’s behavior, and allows any differences between conditions to be attributed to subjects’ concerns over being monitored by others (Reis, 1981). Hypothesis 2 predicts that when the choice process is observed by others and procedural norms can be enforced, group members will create lotteries that are higher in procedural fairness than when the process is not observable. By making higher joint payoffs available to those who are willing to tolerate higher distributional unfairness, the study also tested Hypothesis 3, which predicts that subjects will be less concerned with distributional fairness when the choice process is public than when the choice process is private.

To provide additional evidence that the effect of observability drives differences between conditions, as predicted by Hypothesis 2, the ambiguity of preferences for the choice alternatives was also varied. Subjects made two choices, one in which they chose movies (ambiguous preferences) and a second in which they chose distributions of lottery tickets among the three members of their group (clear preferences). Though tradeoffs among group members’ outcomes are more explicit for lottery tickets than for movie tickets, the distributional implications of choosing a movie can be conceptualized similarly because choosing any one movie is equivalent to choosing a distribution of preference ratings across the three group members. Thus, this manipulation was used to test Hypothesis 4, which predicts that the effect of procedural norms should be stronger when preferences are clear than when preferences are ambiguous.
Subjects

Groups of three friends signed up together to participate in the study. Intact groups were recruited because, as discussed earlier, the importance of process is likely to be highest when group members anticipate frequent future interaction. Moreover, members of intact groups are likely to care about others’ preferences for the outcomes. Thus, to the extent that procedural fairness is higher when the process is observable, the change in behavior can be attributed to a desire to comply with procedural norms.

One hundred and two graduate and undergraduate students formed self-selected groups of three members before signing up to participate in the study. None of the subjects had participated in Study 1. The study was conducted in seven sessions ranging from two to eight groups of participants each. As a reward for participating, subjects were given passes to see the movie selected by their group and state lottery tickets (scratch-off tickets in denominations of $1, $2, $3 and $5) based on the distribution selected by their group.

Stimuli

To select their group’s distribution of lottery tickets, subjects were given a list of twelve distributions of tickets among the three members of their group (see Appendix B). The lottery ticket distributions were designed so that no single distribution gave equal payoffs to all three group members. However, subjects could implement a process giving each group member equal expected outcomes (achieving ex ante fairness) by selecting multiple alternatives, each of which had an equal probability of being selected. The distributions varied in both total joint payoffs and ex post fairness: nine distributions with $9 joint payoffs resulted in smaller differences in payoffs across group members, while three distributions with $10 joint payoffs resulted in larger differences in payoffs across group members.

To select the movie that their group would see together, subjects were given a list of twelve movies currently playing at local theatres (see Appendix B). The movies were selected to represent a variety of genres and ratings, to maximize the variation in subjects’ preferences for the movies. To minimize the possibility that the subjects had already seen the movies, the movies selected were recent releases, and subjects were asked to indicate whether they had seen or had heard about each of the movies.
Design

A major challenge in studying choices made by intact groups is that groups vary dramatically in their histories and norms. One way to control for these differences is by using a within subjects design, in which each group member gives a series of responses under different conditions. For this reason, the design was a 2 x 2 x 2 x 2 mixed factorial design, in which the choice process (Public Choice Process, Private Choice Process) and choice type (Movies, Lottery Tickets) were varied within subjects, and the orders of the choice processes and choice types were counterbalanced between subjects. For both movies and lottery ticket distributions, subjects created lotteries from which the final choices were made in two different conditions, the Public condition and the Private condition. In total, each subject created four lotteries.

Procedure

As groups arrived at the sessions, the group members were seated one behind another, in a row, to minimize their interaction during the first part of the session. They learned that there were two different choice processes their groups would use during the session, the Public Choice Process and the Private Choice Process. They would use one of the processes to choose their lottery tickets and the other process to choose their movie.

For the Private Choice Process, each subject created a private lottery from which the final choice would be made. Each of the three group members had a 1 in 3 chance to provide the lottery for the final choice. To create this lottery, subjects checked off the alternatives they wanted to include in the lottery in the leftmost column of the list of movies or lottery tickets (see Appendix B). Subjects could include as many or as few alternatives as they liked, and each of the alternatives included in this lottery had an equal chance of being selected randomly as the final choice for the group. For example, subjects might create either the lottery \{(5,5,0), (5,0,5), (0,5,5)\} or the lottery \{(5,2,2)\} for the lottery ticket choice, and either the lottery \{EdTV, The Matrix, True Crime\} or the lottery \{EdTV, The Mod Squad\} for the movie choice.

When the Private Choice Process was used, subjects knew that the other group members would not see the lottery they had created; if their lottery were used to select the final choice, others would see only the outcome. In contrast, when the Public Choice Process was used, subjects knew that the other group members would see the lottery they had created (otherwise, the process was the same in for the Private Choice Process). Thus, when the Public Choice Process was used, subjects knew that others would be able to tell whether they had been given an
equal opportunity to receive a preferred outcome – but had been unlucky – or whether the subject had stacked the deck in his or her own favor.

Creating these lotteries allowed subjects to determine both the level of distributional fairness (e.g., including (5,2,2) or (5,5,0)) and the level of procedural fairness of their choice (e.g., including (5,2,2) or {(5,2,2), (2,5,2), (2,2,5)}). Given the alternatives available to the subjects, lower ex ante unfairness could be coupled with either higher or lower ex post unfairness. For lottery tickets, for example, either a combination with low ex post unfairness such as {(5,2,2), (2,5,2), (2,2,5)} or a combination with high ex post unfairness such as {(5,5,0), (5,0,5), (0,5,5)} results in ex ante unfairness of 0.

Subjects were not told initially which process (Public or Private) they would use for each choice. Instead, they were asked to create lotteries for each process for each choice as a contingency. Thus, each subject created four lotteries: a private lottery for movies, a public lottery for movies, a private lottery for lottery tickets, and a public lottery for lottery tickets. This procedure allowed the alternatives selected by each subject for the Public Choice Process to be compared with the alternatives selected by the same subject for the same choice for the Private Choice Process. This within subject comparison controlled for the utility of the outcomes of the choice, because each subject might have different preferences for the alternatives and different relative concerns for their own and others outcomes. This comparison, then, isolated the effects of observability and ambiguity of preferences.

After they had created their lotteries, subjects rated their preferences for each of the twelve movies available to them, and then they rated their perceptions of the other two group members’ preferences for the same movies. These preference ratings were used to compute the subject’s own expected value for the lottery they created for movies, the expected value they predicted for each of the other group members, and the ex post and ex ante unfairness of each lottery. The subjects also answered several questions about their relationships with the other members of their group, such as their frequency of past interaction and predicted future interaction.

Finally, the group learned which of the two choice processes they would use for each choice, and which of the three group members would provide the lottery for each choice. The lottery that would be used for the Public Choice Process was distributed to all three group members for their review. At the end of the session, movie and lottery tickets were distributed to
the subjects. One of the alternatives from the lottery of movies was selected randomly to be the final choice for the group, and subjects were given tickets to see this movie together. One of the lottery ticket distributions was selected randomly to be the final choice for the group, and instant lottery tickets in the appropriate denominations were distributed to the group members.

Results

The first step of analyzing the results was to calculate the subject’s own expected outcome, others’ expected outcomes, ex ante unfairness and ex post unfairness for each of the lotteries that had been created by the subjects. For lottery tickets, expected outcomes were computed using the face values of the lottery tickets included in each distribution. For movies, expected outcomes were computed using the subject’s own ratings of each group member’s enjoyment of the movies. Recall that each subject had rated preferences for each movie from the point of view of each of the group members, implicitly indicating the utility tradeoffs perceived for each movie among the three group members. Thus, both ex ante unfairness and ex post unfairness could be computed as perceived by each subject, so that interpersonal comparisons of utility were limited to those acknowledged by the subjects themselves. Both movie preference ratings and the face values of lottery tickets were normalized prior to computing these measures.

Subjects had been allowed to submit either the same or different lotteries for the Public and Private Choice Processes. Fifty-five percent of the subjects (56 of 102) proposed a different lottery for movies in the Public and Private conditions. Fifty-one percent of the subjects (52 of 102) proposed a different lottery for lottery ticket distributions in the Public and Private conditions. There was no difference in the predicted frequency of future interaction (p > .54) or the frequency of past interaction with other group members (p > .31) for subjects who proposed the same or different menus. There were also no differences in subjects’ beliefs that they knew others’ preferences (p > .69), that they agreed on the activities they enjoyed (p > .63), or that the other group members would consider their preferences when creating their lotteries (p > .48).

However, based on the lotteries that subjects created in the Private condition, it is possible to predict with significant accuracy which subjects had changed their lotteries and which subjects had not. A maximum likelihood probit estimation using the difference between one’s own and others’ expected payoffs, ex ante unfairness and ex post unfairness as the independent variables significantly predicted whether subjects had changed their lotteries across conditions, $\chi^2(3) = 33.47, p < .001$. Those subjects whose lotteries for the Private Choice
Process had the greatest difference between their own and others’ expected outcomes, the highest ex ante unfairness and the lowest ex post unfairness were most likely to change their lotteries across conditions.

*Observability.* Manipulating the observability of the process allowed both Hypothesis 2 and Hypothesis 3 to be tested. Hypothesis 2 predicted that when subjects expected others to see their lotteries, they would create lotteries were more ex ante fair than when they did not expect others to see their lotteries. Hypothesis 3 predicted that subjects would create lotteries that were less ex post fair when they could demonstrate to others that they had followed a fair choice process. Results were analyzed separately for movies and lottery tickets. For both movies and lottery tickets, results were analyzed using a repeated measures GLM in which the choice process (Public, Private) was a repeated factor and the order of choices (Movie choice first, Lottery ticket choice first) and the order of processes (Public first, Private first) were between-subjects factors.

For lottery tickets, ex ante unfairness was significantly higher in the Private condition ($M = .69$) than in the Public condition ($M = .43, F(1,98) = 7.10, p < .01$), supporting Hypothesis 2 (see Figure 3). The between-subjects order of choices, order of processes and their interaction were significant for ex ante unfairness; however, no interactions with the treatment condition were significant. In contrast, ex post unfairness was marginally lower in the Private condition ($M = 2.23$) than in the Public condition ($M = 2.45, F(1,98) = 3.46, p < .10$, providing marginal support for Hypothesis 3. No other effects were significant for ex post unfairness ($ps > .11$).

Thus, while the expected outcomes of the lotteries were more fair in the Public condition, the actual outcomes of the random final choice were on average *less* fair in the Public condition. In the Public condition, subjects were willing to propose more distributionally unfair alternatives than in the Private condition. For example, 23% of the subjects in the Public condition but only 11% of the subjects in the Private condition created the lottery \{(5,5,0)\, (5,0,5),\, (0,5,5)\}. This combination gave the group the maximum joint payoff ($10$ rather than $9$) and gave each subject an equal expected payoff, but resulted in the maximum level of ex post unfairness.

---

4 Ex ante unfairness was higher when the Public process preceded the Private process and when the movie choice preceded the lottery ticket choice, and the latter effect was larger when the Public process was first.
The same pattern of effects was replicated for movies, suggesting that the results are robust. Ex ante unfairness was marginally higher in the Private condition ($M = .29$) than in the Public condition ($M = .26$, $F(1,98) = 2.95, p < .10$), marginally supporting Hypothesis 2 (see Figure 4). No other effects were significant ($ps > .16$). In contrast, ex post unfairness was marginally lower in the Private condition ($M = .48$) than in the Public condition ($M = .52$, $F(1,98) = 3.53, p < .10$, marginally supporting Hypothesis 3. No other effects were significant ($ps > .21$).

Next, results were analyzed using a repeated measures GLM in which both the choice process (Public, Private) and the fairness measure (ex ante, ex post) were repeated factors and the order of choices (Movie choice first, Lottery ticket choice first) and the order of processes (Public first, Private first) were between-subjects factors. This analysis shows a significant interaction between observability and the fairness measure for both movies, $F(1,98) = 13.57, p < .001$ and lottery tickets, $F(1,98) = 11.65, p < .001$, demonstrating that observability had the opposite effect on ex ante and ex post unfairness. For both movies and lottery tickets, higher observability resulted in lower ex ante unfairness but higher ex post unfairness.

Ex post unfairness was significantly higher than ex ante fairness for both movies, $F(1,98) = 48.70, p < .001$ and lottery tickets, $F(1,98) = 145.88, p < .001$, reflecting differences in the way the two measures are calculated. No other effects were significant, other than a 3-way
interaction between the choice process, the order of choice processes and the order of choices, 
\( F(1,98) = 4.02, p = .05 \) for movies, which is of no theoretical interest.

Figure 4: Ex Ante and Ex Post Unfairness of Lotteries for Movies

![Graph showing unfairness of outcomes for movies.](image)

Figure 5: Comparison of Lotteries Changed/Not Changed Across Conditions (Lottery Tickets)

![Graph showing expected payoffs for lottery tickets.](image)

Why did subjects create different lotteries in the Public and Private conditions? A more
detailed analysis of the data suggests that subjects who modified their lotteries across conditions
did so in a manner that reflected social desirability concerns. For example, for lottery tickets, the
fifty-two subjects who changed their lotteries across conditions created lotteries that gave
themselves significantly higher expected payoffs \( (M = .48) \) than the fifty subjects who did not
change their lotteries \( (M = .04; t(100) = 4.57, p < .001) \) in the Private condition. While those
who did not change their lotteries across conditions gave themselves and others the same
expected payoffs ($p > .29$), those who changed their lotteries gave themselves significantly higher payoffs ($M = .48$) than they gave to others ($M = -.21$, paired $t(51) = 5.07, p < .001$) in the Private condition and they gave themselves marginally higher payoffs ($M = .16$) than they gave to others ($M = -.04$; paired $t(51) = 1.72, p < .10$) in the Public condition (see Figure 5).

In summary, as predicted by Hypotheses 2 and 3, the relative importance of distributional fairness and procedural fairness changes when the choice process (the lottery) can be observed by others. When ex ante fairness can be demonstrated to others, as in the Public condition, ex post unfairness can be more clearly attributed to situational variables (e.g., random chance) rather than to the group member who created the lottery. In contrast, when others can observe only the outcome, as in the Private condition, subjects seemed to give more weight to ex post fairness. This difference across conditions is consistent with Blount’s (1995) demonstration that preferences for fair outcomes are affected by beliefs about how the outcomes transpired (e.g., whether unfair outcomes were selected by a human or by a computer). The interesting finding in this study is that subjects appeared to anticipate this attribution process.

Ambiguity of Preferences. Further evidence that these effects are driven by observability is provided by comparing the lotteries created for the movie and lottery ticket choices. When preferences are clear (e.g., for lottery tickets, which have a face value), the lotteries that are created can be more easily evaluated as fair or unfair than when preferences are ambiguous (e.g., for movie tickets). Hypothesis 4 predicted that observability would have a stronger influence on the proposer’s behavior when preferences for the alternatives were clear. This study tested the interaction between observability and ambiguity of preferences by comparing lotteries of movies (ambiguous preferences) with lotteries of lottery tickets (clear preferences) when the two choices were made under the same conditions (Public or Private).

The degree of bias towards the subjects’ own payoffs, measured by calculating the difference in expected payoffs between the subject and the two other group members, varied significantly across conditions (see Figure 6). Results were analyzed using a repeated measures GLM in which both the choice process (Public, Private) and the choice (movies, lottery tickets) were repeated factors and the order of choices (Movie choice first, Lottery ticket choice first) and the order of processes (Public first, Private first) were between-subjects factors. The within subject effect of condition was significant, showing that for both choices, bias was smaller in the Public condition ($M = .11$) than in the Private condition ($M = .26$), $F(1,98) = 13.51, p < .001$. 

80
However, this effect was qualified by a significant interaction between the choice process and the choice, supporting Hypothesis 4. The effect of observability was stronger for lottery tickets ($M = .29$ in the Private condition > $M = .04$ in the Public condition) than for movies ($M = .23$ in the Private condition > $M = .18$ in the Public condition), $F(1,98) = 6.29, p = .01$. Only one other effect was significant, the interaction between choice and order of choices, indicating that the order effect was stronger for lottery tickets than for movies, $F(1,98) = 7.44, p < .01$.

*Figure 6: Observability Had More Impact When Preferences Were Clear*

![Graph showing expected value for movies and lottery tickets under private and public conditions.]

When the results for movies were analyzed separately, the effect of the choice process was only marginally significant, $F(1,98) = 2.79, p < .10$. Lotteries for movies gave significantly higher expected payoffs to the subject than to other group members in both the Private condition (paired $t(101) = 3.72, p < .001$) and in the Public condition (paired $t(101) = 2.92, p < .01$). In contrast, when results were analyzed separately for lottery tickets, the effect of choice process was significant, $F(1,98) = 10.99, p < .001$. Lotteries for lottery tickets gave significantly higher payoffs to the subject than to other group members in the Private condition (paired $t(101) = 3.06, p = .01$) but not in the Public condition, ($p > .59$). For lottery tickets, bias was completely eliminated when the choice process was observable.

In summary, as predicted by Hypothesis 4, the effect of observability was stronger when preferences for the alternatives were clear (a choice of lottery tickets) than when preferences...
were ambiguous (a choice of movies). The significant interaction between observability and ambiguity of preferences provides additional evidence that the effect of observability was driven by a desire to comply with procedural norms. When it was more difficult to evaluate compliance, for the movie choice, subjects took advantage of the ambiguity to choose movies they predicted they would like more than others would like. One implication of these results is that clarifying preferences (e.g., by rating one's preferences for the movies being considered and sharing them with the group) may help to achieve fairness in group choices.

Discussion

The results of Study 2 support Hypotheses 2, 3 and 4. Knowing that others will observe an intermediate outcome of the choice process affects not only the procedural fairness of the lotteries created, but also the distributional fairness of the alternatives included in the lottery. Lotteries created for two different choices, movies and lottery tickets, showed a consistent pattern of results for both ex ante unfairness and ex post unfairness measures. Subjects seemed to be more tolerant of distributional unfairness when they could demonstrate to others that they had used a fair process to arrive at the outcomes. Moreover, the finding that observability has stronger effects when preferences are clear than when preferences are ambiguous increases our confidence that the differences between conditions are motivated by a desire to comply with procedural norms.

Building on the results of Study 1, this study demonstrated that people consider ex ante fairness not only when they evaluate lotteries, but when they make choices for others. Though the phased choice process used in this study is somewhat stylized, asking subjects to create lotteries from which choices would be made randomly allowed subjects to determine both the distributional fairness of the choice and the procedural fairness of the choice. This allows a comparison of the relative importance of distributional fairness and procedural fairness across conditions.

In addition to the predicted results, an interesting observation based on the results of Study 2 is that subjects seemed to avoid taking responsibility for making choices that affected others, especially in the Public condition. Despite the fact that all subjects were given the freedom to choose a single movie or single lottery ticket distribution rather than submit multiple solutions.

5 This effect cannot be explained by a difference between the ratings of the subject’s own and others’ preferences. For the full set of 12 available movies, there was not a significant difference between the subject’s own ratings ($M = -.04$) and the subject’s perceptions of others’ ratings ($M = -.04, p > .97$).
alternatives to a lottery, only a small fraction of subjects did so. Thirty-five subjects (34%) chose a single movie in the Private condition, while only nine subjects (9%) chose a single movie in the Public condition. Similarly, only thirteen subjects (13%) in the Private condition and three subjects in the Public condition (3%) chose a single distribution of lottery tickets. The fact that this percentage differs significantly across conditions (by a factor of four) suggests that subjects understood their ability to create single alternative lotteries, but constrained their behavior to comply with social norms.

**General Discussion and Conclusions**

Previous research provides considerable evidence that people care about the process used to make choices in addition to the outcomes. Procedural fairness has been operationalized in many different ways, such as by varying trial procedures (e.g., Latour, 1978; Lind et al., 1993), the accuracy of an assessment test (e.g., van den Bos, Vermunt & Wilke, 1997), the method for computing course grades (e.g., Tyler & Caine, 1981), and allowing or not allowing subjects to voice their opinions (van den Bos, Lind, Vermunt & Wilke, 1997).

This diversity of definitions for procedural fairness is both a strength and a weakness of the literature. It is a strength because it shows that preferences for fair procedures are robust to a variety of operationalizations, encompassing multiple criteria, such as consistency across people, consistency across time, bias suppression, accuracy, correctability, and representativeness (Leventhal et al., 1980). However, this diversity is also a weakness, because previous research has been able to capture these variables only in a binary or categorical manner, finding no single measure to capture all of the operationalizations. By selecting a few of these criteria, consistency across people and bias suppression, this research showed that procedural fairness can be expressed in a more quantitative manner, as a criterion of “equal opportunity.”

The results of two studies demonstrated that the criterion of equal opportunity is not only a normative criterion, but also a descriptive one: people seem to intuitively take probabilities and expected values into consideration when assessing fairness. After introducing a more quantitative, continuous measure of procedural fairness, ex ante fairness, this paper showed that the measure accurately predicted group members’ fairness judgments and preferences for allocations. The results presented in this paper demonstrate that this measure is predictive not only for monetary allocations, such as instant lottery tickets, but also for decisions that are made.
by groups on a more regular basis, such as choosing a movie that group members will watch together.

This paper also contributes to the literature by demonstrating that the decision of procedural fairness can be endogenous to a choice rather than exogenous. Previous research had examined people's assessments of procedural fairness, but had not examined whether people are able to implement a normatively fair choice process, taking a proactive role in defining the process. When procedural fairness is conceptualized in a stylized way, as a lottery, people can be given the opportunity to implement a fair (or unfair) process for allocating resources to themselves and other group members. Consistent with theoretical predictions, Study 2 showed that people implemented a more procedurally fair and less biased process when the choice process could be observed by others and interpreted unambiguously (lottery tickets vs. movies).

Measuring and conceptualizing procedural fairness in a more continuous manner is important because it allows tradeoffs between distributional fairness and procedural fairness to be examined. Distributional fairness can be thought of as a constraint in allocation decisions. As it is defined in this paper, ex ante fairness is a less limiting constraint than ex post fairness because an allocation can be ex ante fair even if it not ex post fair, and an allocation is always at least as fair ex ante as it is ex post. If, as the results of this research suggest, a publicly observable choice process increases the weight on procedural fairness relative to distributional fairness, then efficiency might be increased by using a public choice process rather than a private choice process.

Thus, this research has implications for the balance between equity and efficiency in choices made by groups. The results suggest that the choice process used by a group – public or private – can influence whether more efficient but ex post unfair alternatives are considered by the group. For example, when subjects chose distributions of lottery tickets in Study 2, they were more likely to choose distributions with higher joint payoffs ($10 for the group rather than $9) but lower distributional fairness (two group members received $5 tickets and one group member received nothing) in the public condition than in the private condition. As in this experiment, utility maximization often requires groups to consider ex post unfair alternatives.

Such process considerations are relevant for organizations relying on groups to make choices on their behalf. In the experiments reported here, the observability of the group's choice process was manipulated as an independent variable. In contrast, in natural settings, group
processes may be determined by the group’s norms or by standard organizational procedures. Understanding the role of procedural norms in group choices helps to predict when being able to demonstrate ex ante fairness might encourage group members to consider alternatives that are ex post unfair but provide higher joint utility. From a normative perspective, these results suggest that organizations should carefully consider the implications of the choice process when designing procedures for group choices.
References


Appendix A

To compute ex ante and ex post unfairness, equations proposed by Boiney (1995) were adapted.

First, the unfairness perceived by each group member (“envy”) for each alternative in the lottery is computed as the average of the positive differences between others’ outcomes and the group members’ own outcomes. For example, using the equation below to compute envy for an alternative giving $5 to Member 1, $5 to Member 2 and nothing to Member 3, gives the result that Member 1 and Member 2 have envy of 0, but Member 3’s envy is 5.

\[
e_{i}(x) = \left(\frac{i}{n-1}\right) \sum_{j} k_{j} e_{ij}(x)
\]

where

- \(k_{j}\) is a weight greater than 0 (to simplify, assume \(k_{j}\) is 1),
- \(x\) is the vector of expected consequences for a particular outcome,
- \(e_{ij} = \max(0, u_{i}(x_{j}) - u_{i}(x_{i}))\), a measure of pairwise disadvantageous envy,
- \(n\) is the number of group members.

Ex post unfairness is a measure of the dispersion of individual envy across group members, and it is calculated for each of the possible alternatives that may result from the lottery. The ex post unfairness of this distribution is 16.67, the sum of the squared differences between each member’s envy and the average envy across group members (in this case, 1.67).

\[
I_{p}(x) = \text{ex post unfairness} = \left[\sum_{i} (e_{i}(x) - \bar{e}(x))^{2}\right]/n
\]

where

- \(e_{i}(x)\) is individual envy,
- \(\bar{e}(x)\) is the average envy across group members for outcome \(x\).

For a lottery over alternatives, expected ex post unfairness is calculated as an expectation over each possible final outcome that may result from the lottery. For example, suppose that the lottery includes the alternatives \{A, B, C\}. Ex post unfairness is first calculated separately for the outcomes of A, B, and C, and then expected ex post unfairness is computed as the sum over A, B and C of the probability of each outcome multiplied by the ex post unfairness for each outcome.

---

6 Setting \(k=1\) assumes that the proposer’s outcomes are not perceived differently relative to the other group members’ outcomes. However, it is quite possible that those observing the choice process might evaluate unfairness which is favorable to the proposer differently from unfairness which is unfavorable to the proposer (Blount, 1995). In this case, envy of the proposer should be given greater relative weight.
Ex ante unfairness is computed as the sum of squared differences between each group member’s expected envy and the average expected envy for all three group members. If a random choice is made from each lottery, giving each alternative equal probability, then expected envy can be computed as each group member’s average envy across the alternatives included in the lottery. For example, suppose that three players face the following lottery over outcomes, and that each of the three outcomes has an equal chance of being selected:

Outcome 1: $5 to Player 1, $5 to Player 2, and 0 to Player 3
Outcome 2: $5 to Player 1, 0 to Player 2, and $5 to Player 3
Outcome 3: 0 to Player 1, $5 to Player 2, and $5 to Player 3

In this case, the expected envy for each group member is 1.67, average envy is 1.67, and, using the equation below, ex ante unfairness is 0.

\[ I_A = \text{ex ante unfairness} = \sum_a \left[ E(e_i(a)) - E(\overline{e(a)}) \right]^2 / n \]

where
- \( a \) is the lottery
- \( e_i(a) \) is expected individual envy for the lottery
- \( E(\overline{e(a)}) = \sum_{x \in a} p(x) e(x) \) is expected average envy across group members
- \( p(x) \) is the probability of outcome \( x \)
## Appendix B

<table>
<thead>
<tr>
<th>Include</th>
<th>Movie Title</th>
<th>Type of Film</th>
<th>Stars</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Analyze This</td>
<td>Comedy</td>
<td>Billy Crystal, Robert De Niro, Lisa Kudrow</td>
<td>R</td>
</tr>
<tr>
<td></td>
<td>Baby Geniuses</td>
<td>Family</td>
<td>Kathleen Turner, Christopher Lloyd, Kim Cattrall, Peter MacNicol, Ruby Dee</td>
<td>PG</td>
</tr>
<tr>
<td></td>
<td>Cruel Intentions</td>
<td>Drama</td>
<td>Sarah Michelle Gellar, Ryan Phillippe, Reese Witherspoon, Selma Blair</td>
<td>R</td>
</tr>
<tr>
<td></td>
<td>EdTV</td>
<td>Comedy</td>
<td>Matthew McConaughey, Jenna Elfman, Woody Harrelson, Ellen DeGeneres</td>
<td>PG-13</td>
</tr>
<tr>
<td></td>
<td>Forces of Nature</td>
<td>Romantic Comedy</td>
<td>Ben Affleck, Sandra Bullock, David Strickland, Blythe Danner, Joe Don Baker, Steve Zahn, Meredith Scott Lynn</td>
<td>PG-13</td>
</tr>
<tr>
<td></td>
<td>The King and I</td>
<td>Family</td>
<td>Miranda Richardson, Martin Vidnovic, Christiane Noll, Ian Richardson, Darryl Hammond</td>
<td>G</td>
</tr>
<tr>
<td></td>
<td>Lock Stock and Two Smoking Barrels</td>
<td>Comedy</td>
<td>Nick Moran, P.H. Moriarty, Sting, Dexter Fletcher, Jason Statham, Jason Flemyng, Steven Macintosh</td>
<td>R</td>
</tr>
<tr>
<td></td>
<td>The Matrix</td>
<td>Action/Adventure</td>
<td>Keanu Reeves, Laurence Fishburne, Carrie-Anne Moss, Joe Pantoliano, Julian Arahanga</td>
<td>R</td>
</tr>
<tr>
<td></td>
<td>The Mod Squad</td>
<td>Action/Adventure</td>
<td>Claire Danes, Omar Epps, Giovanni Ribisi, Dennis Farina</td>
<td>R</td>
</tr>
<tr>
<td></td>
<td>Ten Things I Hate About You</td>
<td>Romantic Comedy</td>
<td>Heath Ledger, Julia Stiles, Joseph Gordon-Levitt, Andrew Keegan, Larisa Oleynik</td>
<td>PG-13</td>
</tr>
<tr>
<td></td>
<td>Wing Commander</td>
<td>Action/Adventure</td>
<td>Freddie Prinze Jr., Matthew Lillard, Saffron Burrows, Jurgen Prochnow</td>
<td>PG-13</td>
</tr>
<tr>
<td>Include (✓)</td>
<td>Distribution</td>
<td>Member 1's Ticket</td>
<td>Member 2's Ticket</td>
<td>Member 3's Ticket</td>
</tr>
<tr>
<td>------------</td>
<td>--------------</td>
<td>-------------------</td>
<td>-------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>A</td>
<td>$3 ticket</td>
<td>$5 ticket</td>
<td>$1 ticket</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>$5 ticket</td>
<td>$1 ticket</td>
<td>$3 ticket</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>$2 ticket</td>
<td>$2 ticket</td>
<td>$5 ticket</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>$1 ticket</td>
<td>$3 ticket</td>
<td>$5 ticket</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>$5 ticket</td>
<td>$2 ticket</td>
<td>$2 ticket</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>$5 ticket</td>
<td>$5 ticket</td>
<td>no ticket</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>$3 ticket</td>
<td>$1 ticket</td>
<td>$5 ticket</td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>no ticket</td>
<td>$5 ticket</td>
<td>$5 ticket</td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>$1 ticket</td>
<td>$5 ticket</td>
<td>$3 ticket</td>
<td></td>
</tr>
<tr>
<td>J</td>
<td>$5 ticket</td>
<td>$3 ticket</td>
<td>$1 ticket</td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>$2 ticket</td>
<td>$5 ticket</td>
<td>$2 ticket</td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>$5 ticket</td>
<td>no ticket</td>
<td>$5 ticket</td>
<td></td>
</tr>
</tbody>
</table>
Enduring Effects of Social Context on Categorizing Customer Needs*

* This paper is co-authored with Nader T. Tavassoli, MIT Sloan School of Management. Work reported here was supported by the MIT Center for Innovation in Product Development under NSF Cooperative Agreement Number EEC-9529140. The authors would like to thank David Bell, John Carroll, John Hauser, Drazen Prelec and JoAnne Yates for their helpful comments, and Debbie Shoap and Drazen Prelec for collaborating with us on the marketing research projects reported here.
Enduring Effects of Social Context
on Categorizing Customer Needs

Abstract

Structured categorization techniques are used in market research and product development to better understand underlying customer needs for products and services. The results produced by these techniques affect strategic decisions such as product positioning and product development. Some techniques are used primarily by individuals and others with groups, but often the social context in which they are used is considered interchangeable. However, both social factors (such as anticipated evaluation by others) and procedural factors (such as the decision rules used by the group) suggest that whether these techniques are used individually or by a group may significantly change the results. Differences in the way customer needs are categorized – into broader or narrower groupings, for example – are important because the next step of the process might be to map these higher level customer needs to product attributes or to define an advertising strategy that will address these needs.

In two studies, we observe qualitative differences in the way individuals and groups categorize the same set of customer needs. Moreover, under certain conditions, whether a categorization technique is learned in an individual or group setting can have a lasting effect on the way the technique is used, even in a different social context. In a third study, we examine one of the potential mechanisms causing the enduring effects we observe. These systematic differences and their enduring effects represent an opportunity for managers, because they can be leveraged when designing training strategies or sequencing activities in the product development process. Recognizing that the categorization structures created are only one of the issues important to managers, we also discuss organizational considerations such as consensus among group members, the satisfaction of participants with the outcomes, and the comprehensiveness of the customer needs considered.
Introduction

Categorization is central to marketing decision making, both from the consumer's perspective and from the firm's perspective. Because categories shape consumers' perceptions of products and determine the alternatives to which a given product is compared, the way consumers categorize information can affect their choices (e.g., Urban, Hulland & Weinberg, 1993; Sujan & Bettman, 1989). From the firm's perspective, understanding how consumers categorize is an important input into marketing strategy and product positioning (e.g., Hauser, 1986; Ries & Trout, 1986). Because structured categorization techniques are widely used in the product development process, categorization behavior often has a direct influence on product design (e.g., Burchill & Brodie, 1997; Griffin & Hauser, 1993; Hauser & Clausing, 1988; Shiba, Graham & Walden, 1993).

We examine categorization behavior in the context of product development. While a variety of techniques for analyzing customer needs have been designed for use by individuals and groups, we argue that whether a structured categorization technique is used individually or by a group qualitatively influences the way concepts are categorized. Moreover, we suggest that these qualitative differences may be enduring, influencing not only current performance, but in some cases individuals' subsequent performance using the same technique. Systematic differences in the learning and use of such techniques by groups and individuals represent an opportunity for managers because these differences can be leveraged when designing training strategies or sequencing activities in the product development process.

We compare categorization structures created by individuals and groups using two techniques widely adopted in industry. First, we examine a relatively simple technique called the customer sort in which customer needs for a product are divided into logically related piles (see Griffin & Hauser, 1993). The customer sort technique (e.g., AMS's Vocalyst™ method) is designed to collect and statistically analyze large numbers of observations from individuals (and occasionally groups) who categorize the same set of customer needs. Next, we examine a more complex technique, called an affinity diagram, in which customer requirements are combined into hierarchical categories on a poster-sized sheet of paper (see Burchill & Brodie, 1997; Shiba et al, 1993). Affinity diagrams, also known as LP Diagrams or KJ Diagrams, are used primarily by teams but are also recommended for use by individuals. In previous research, Griffin and Hauser (1993) compared the effectiveness of affinity diagrams created by groups with customer
sorts completed by individuals, but their analysis did not isolate the effect of social context. In contrast, we specifically examine the effect of social context (working individually or working as part of a group) on the use of these techniques.

The assumption underlying the use of such techniques is that the relationships participants identify among concepts reveal their tacit knowledge about the concepts. Consider, for example, a customer who needs a vehicle to drive to work. Among dozens of needs being categorized, the customer might combine the phrases “spacious headroom” and “provides lumbar support” into a category described as “I feel comfortable in the driver’s seat.” The same customer might combine “high quality stereo sound” and “engine runs quietly” into a second category based on the notion that “I want a great sound system.” If using a hierarchical categorization technique, these two categories might be combined into the higher level category “I want a comfortable interior.” In contrast, a different customer categorizing the same set of requirements might combine “engine runs quietly” with “vehicle can run for two years before the first tune-up,” because both suggest low-effort maintenance. This difference in the way the two customers categorize their needs might suggest to product designers that there are at least two customer segments with different concerns.

Our research makes two contributions. First, we examine the effect of social context on categorization behavior. In Study 1, we compare groups and individuals using a single-level categorization technique, the customer sort. In Study 2, we examine these differences for a hierarchical categorization technique, the affinity diagram. Second, we examine how an initial group experience affects the subsequent individual use of the same technique for different information, and conversely, how an initial individual experience affects the subsequent use of the technique in a group setting. Finally, in Study 3, we investigate one of the potential mechanisms for the enduring effects we observe in Study 2. We examine whether seeing a visual example of information categorized for an unrelated topic can have an effect on categorization behavior similar to that of prior experience.

In the next section, we explain why we expect social context to influence categorization behavior. Next, we outline the conditions under which the effects of social context may be enduring. We guide the reader through three studies testing our hypotheses. Finally, we conclude by discussing the implications of our findings for organizations using structured categorization techniques, including considerations such as consensus among participants,
coverage of the information domain, and the participants’ satisfaction with the output.

Social Context

While categories may be generally agreed upon when items have similar attributes or features, there are often multiple ways to categorize a set of items. Early research in cognitive psychology suggested that categorization behavior was a personality trait – that some people categorize more narrowly than others, regardless of the domain (Gardner, 1953; Gardner & Schoen, 1962). However, subsequent research has shown that categorization behavior is somewhat malleable. For example, the knowledge and experience of the categorizer can influence the way items are categorized. Experts tend to differentiate more among concepts than novices (Johnson & Mervis, 1997). One’s affective state also may influence the way concepts are categorized. Subjects in whom positive affect had been induced categorized more inclusively than control subjects (Isen & Daubman, 1984). Our paper adds to this literature by demonstrating that social context, working individually or as part of a group, may systematically influence categorization behavior.

Individuals often enter a group with divergent perspectives on the objects being categorized. In fact, in the case of cross-functional teams, groups may be constructed specifically because they have divergent perspectives. We suggest that the process of integrating individual views will cause systematic differences between group and individual categorization judgments. Both social factors and procedural factors may contribute to these differences. For example, group members may feel social pressure to state opinions evaluated favorably by others, or they may be convinced by information that is exchanged during group discussion (Laughlin & Earley, 1982). Procedural factors may limit the degree of integration that occurs. For example, there may be limits to the number of items that a group can comfortably consider simultaneously (Hinsz, Tindale & Vollrath, 1997). Conversational norms within groups may limit rapid changes in topics, a phenomenon that has been called the “focus effect” (Delbecq, van de Ven & Gustafson, 1975). These factors all affect the shared understanding that develops within groups.

Just as individuals categorize based on their mental models of the relationships among concepts (Medin, Lynch, Coley & Atran, 1997), a categorization structure created by a group reflects a shared understanding of the relationships between concepts (Massey & Wallace, 1996). A mental representation derived by a group is, essentially, the result of individuals coming to
consensus over how to integrate their individual mental representations (Carley, 1986). However, this aggregation process is unlikely to be a simple averaging of individual perspectives. Instead, Walsh, Henderson & Deighton (1988) suggest that the structure of the group’s collective map is a ‘negotiated belief structure,’ influenced by political processes within the group.

The mental representations of groups might be either more, equally, or less complex than the mental representations of individuals (for a review, see Hinsz et al., 1997). Based on theoretical arguments, it has been suggested that groups perceive fewer dimensions in complex stimuli than individuals (Doise, 1978). However, a test of this idea using multidimensional scaling to analyze similarity ratings provided by groups and individuals suggested that they based their ratings on the same number of underlying dimensions (Hinsz, Vollrath, Nagao & Davis, 1988). The opposite view, that members of interactive groups generate more complex mental maps, has also been proposed, and empirical research addressing these conflicting views has been inconclusive (Hinsz et al., 1997). One reason may be that the hierarchical structure has typically been inferred by statistically analyzing inputs such as pairwise similarity ratings. We argue that the process of rating the similarity between items may be quite different from creating an explicit categorization structure.

Organizing information by categorizing or sorting ideas may be especially helpful for uncovering unexpressed assumptions and transmitting knowledge when the problem solving is done in a group setting (Klimoski & Mohammed, 1994; Massey & Wallace, 1996). Individuals categorizing items have their own ideas about the similarity between items, even defining categories such as “things that don’t go with anything else” (Gardner & Schoen, 1962). However, when groups create a shared category structure, multiple group members must be convinced of the relationships among items, and categories cannot be based on connections that are only meaningful to one of the group members.

Thus, one reason similarity judgments differ from explicit categorization is that putting multiple items into a category requires shared beliefs about the similarity of multiple items. To maintain a category’s identity, each item must be similar to several other items simultaneously, and the dimensions of similarity must be agreed upon by all of the group members. This becomes more difficult as the size of the category grows. The smaller the number of items in a category, the greater the probability that individual group members agree that the items are
similar yet have different beliefs about what draws the items together. However, more and more assumptions must be made explicit rather than tacit as the group considers adding additional items to a category. The very strength of techniques like affinity diagrams in bringing out tacit knowledge convinces us that groups will differentiate more among concepts than individuals.

Together, these procedural factors—limits on the number of items a group can entertain simultaneously, the focus effect, and the process of negotiating a mental model based on shared beliefs about the relationships among concepts—suggest that groups will categorize less inclusively than individuals. Given the same number of concepts, groups should create a larger number of smaller categories than individuals. More formally,

*Hypothesis 1:* Individuals will categorize concepts more inclusively than groups, who will differentiate more among concepts.

There are three additional conceptual implications of this hypothesis which apply to hierarchical categorization techniques such as affinity diagrams (see Figure 1). First, if the technique allows participants the freedom to decide whether or not to use each of the available customer requirements, as in the case of the affinity diagram, Hypothesis 1 suggests that groups will use fewer of the available requirements than individuals. Just as individuals will be more inclusive when creating each category, they will be more inclusive when determining whether each idea is relevant to the topic being addressed by the diagram.

*Figure 1: Two Examples of Affinity Diagrams*

![Diagram A and Diagram B](image)

Each diagram represents a poster-sized sheet of paper with the requirements participants used (shaded rectangles) and the various levels of categories they identified (boundaries drawn around the requirements).
For hierarchical categorizations such as affinity diagrams, Hypothesis 1 predicts that at higher levels of abstraction (i.e., categories of categories) there should be more integration in individual diagrams than in group diagrams. Consistent with their behavior in grouping together more requirements at the first level of abstraction, individuals should be more inclusive in grouping together categories at higher levels of abstraction. We will refer to the total number of interactions defined among concepts as diagram complexity. In addition to being theoretically grounded, this prediction is consistent with anecdotal evidence suggesting that teams work better with categorization structures that make clear separations among concepts. For example, while studying teams creating affinity diagrams, Griffin (1989) observed one team decide to eliminate a high level category that combined several concepts in favor of letting each subcategory stand on its own.

Finally, the degree of diagram organization reflects whether interactions among elements are perceived to have been determined using clear rules (Volkema, 1988), and we will refer to this measure as diagram structure. If individuals are more integrative in grouping together concepts and identify more interactions among concepts, individual diagrams may appear to have more inherent structure than group diagrams. Thus, based on Hypothesis 1, we predict that the structure of individual affinity diagrams will be given higher ratings by experts (e.g., users of the data) than the structure of group affinity diagrams. The implications of this prediction are important: if diagrams that are perceived to have more structure more clearly communicate the results, it is possible that a more structured diagram will have more impact on decisions made based on the results.

In summary, Hypothesis 1 predicts a consistent pattern of effects for several dependent measures. A higher degree of concept differentiation by groups should be reflected in a smaller number of customer requirements perceived to be relevant to the topic being addressed, in a smaller number of requirements included in each category, and in a smaller number of interactions defined among categories. Moreover, a group’s tendency to be less integrative in defining relationships among concepts suggests that others will perceive their category hierarchies to be less structured.

**Enduring Effects**

If social context influences the way information is categorized, how might the social context of one’s initial experience using a technique influence future performance? To the extent
that the social context affects the way a categorization technique is used initially, it also affects how the technique is learned. Research on anchoring, goal setting, and habit formation suggests that there may be lasting effects when a person learns to use a categorization technique either individually or in a group setting.

The concept of 'transfer' between tasks performed in sequence has been investigated in the organizational training literature (for a review, see Goldstein, 1993). Completing a related task prior to the experimental task, whether or not participants are aware of any relationship between the two tasks, has been shown to influence both task planning and performance (Earley & Perry, 1987). For example, individuals who completed a hierarchical priming task, first listing categories and then listing ideas within each category, were more likely to use a hierarchical strategy for a subsequent idea generation task than individuals who had not been primed (Earley & Perry, 1987). More specific to categorization behavior, priming tasks have been demonstrated to affect concept differentiation (Lindman, 1976). After subjects had completed a priming task to identify similarities between items, they identified fewer categories from the same number of objects (lower concept differentiation) than the control group. After subjects had completed a priming task to identify differences between items, they identified more categories from the same number of objects (higher concept differentiation) than the control group.

Research on habitual routines suggests that carryover effects may occur at both individual (Weiss & Ilgen, 1985) and group levels of analysis (Gersick & Hackman, 1990). Learning can also be transferred from individual to group settings, and vice versa, a phenomenon which has been referred to as "importation" (Gersick & Hackman, 1990). For example, previous research has demonstrated improvements in group performance due to individual training rather than group training (e.g., Bottger & Yetton, 1987).

Both groups and individuals may be susceptible to passive anchoring effects, insufficient adjustment of judgments from an initial starting point (Whyte & Sebenius, 1997; Rutledge, 1993). Anchoring effects are potentially important because research on goal setting in both individual and group settings indicates that one's expectations about performance may influence one's actual performance on a task (Hinsz & Matz, 1997). Expectations for future performance may be influenced by factors such as the availability of feedback or visibility of output from a prior task (Hackman & Morris, 1975). Thus, transfer across individual and group contexts may be especially strong when the same technique is used in both contexts, because situational cues
are some of the most important factors influencing whether transfer occurs (Goldstein, 1993).

Even if not physically present, memory for visual artifacts created during the performance of the first task, such as affinity diagrams, may provide mental models that shape expectations for later tasks. Memory of the visible output from a previous task is likely to have a particularly strong impact because individuals may adjust their performance in response to external information about appropriate targets (Harkins & Szymanski, 1988). When there are strong cues provided by visual output from the first task, a high degree of consistency may evolve spontaneously yet systematically across subsequent applications of a technique, even if the technique is not used in the same social context. Thus, we predict enduring effects of social context for the more visual dimensions of the affinity diagrams, such as the number of requirements used, the size of the categories created, the complexity of the diagram, and the overall structure of the diagram.

Hypothesis 2: When the social context shapes the visual cues that result from using a technique, the social context in which the categorization technique is learned will have a lasting effect on the way the categorization technique is used, even in a different social context.

To test our hypotheses, we examine two different categorization techniques in a product development setting, the customer sort and the affinity diagram. Study 1 examines the effect of social context on using the customer sort technique in either a group or individual setting. Study 2 examines the effect of social context on hierarchical categorization using affinity diagrams, and tests our predictions about the enduring effects of social context. Finally, Study 3 examines a potential mechanism for the enduring effects that are observed in Study 2.

Study 1

The goal of Study 1 was to test the first hypothesis that individuals and groups will categorize the same set of customer needs in a systematically different manner. Groups should differentiate more among the customer needs than individuals, creating smaller categories. Individuals should be more inclusive in categorizing the same customer needs, creating larger categories.

Experimental Task

The customer sort technique involves the quantitative analysis of categorizations made by many individual customers. The first step is to conduct one-on-one customer interviews, which are transcribed and analyzed to extract the key concepts and phrases that become the customer
needs. These customer needs are printed onto decks of index cards for other customers to sort. Individual customers then sort the cards into piles that are meaningful to them.

For this study, we used a subset of 48 representative customer requirements selected from 220 requirements for food storage containers (i.e., picnic baskets) collected by Griffin (1989). Two criteria were used to select this subset of customer requirements: first, representativeness and second, relatively low co-occurrence in Griffin’s analysis (each possible pairing of the requirements we chose had been put into the same pile by fewer than 35 of Griffin’s 60 respondents). This screening process eliminated the more obvious pairings for which there would be little variance among the subjects.

Subjects

Sixty-one graduate and undergraduate student subjects participated in the experiment, and were given a movie ticket for their participation. Forty-four subjects were assigned to eleven groups of four subjects per group, and seventeen subjects participated individually.

Procedure

Subjects were asked to sort customer requirements for food storage containers as part of a product development effort. They were given instructions for the task (based on instructions used by Griffin, 1989) and an envelope of cards pre-printed with the customer requirements.

In the individual condition, subjects worked by themselves to sort the cards into categories that were logical to them. They were instructed to create as many or as few categories as they felt were necessary. After they had divided the requirements into categories, they selected an exemplar for each category (the requirement they felt was most representative of the category), and placed it on top of the pile of cards. This step minimizes the tendency to create a ‘miscellaneous’ pile composed of requirements only related by the fact that they do not relate to other categories (Gardner & Schoen, 1962). Subjects paper clipped each pile together before putting the piles back into the envelope. In the group condition, the process was the same except that the four subjects assigned to each group worked together to sort the requirements. Each group was given only one set of cards.

After they had finished sorting the requirements, subjects in both conditions answered several questions about their satisfaction with the task and with the outcome. The subjects also documented the number of minutes they took to complete their sort.
Results

There were no differences between conditions in subjects’ familiarity with the product category being evaluated, as measured by their responses to the question “Have you been on a picnic within the last year?” Ninety-four percent of those in the individual condition and 91% of those in the group condition responded affirmatively, $p > .70$. There were no differences between conditions in subjects’ satisfaction with the process (using a 1-7 scale where 7 indicates the highest level of satisfaction, $M = 5.73$ in the group condition vs. $M = 5.47$ in the individual condition; $p > .54$) or satisfaction with the outcome (using a 1-7 scale, $M = 5.59$ in the group condition vs. $M = 5.53$ in the individual condition; $p > .86$). Subjects in both conditions spent the same amount of time on the task ($M = 20.45$ minutes in the group condition vs. $M = 19.63$ minutes in the individual condition; $p > .69$).

The primary dependent variable we consider is the degree of concept differentiation. The customer sort technique is very similar in structure to the Object Sorting Test developed by Gardner (1953) to test of concept differentiation. Thus, the measure of concept differentiation we use is the equivalence range measure developed by Gardner, which is calculated by dividing the number of customer needs (in this case, 48) by the number of categories created, giving the average number of customer needs included in each category.

Groups sorted the 48 requirements into more different categories ($M = 12.55$) than individuals ($M = 10.06$; $t(26) = 2.14, p = .04$). Thus, our measure of concept differentiation, equivalence range, was significantly smaller for groups ($M = 4.05$) than for individuals ($M = 5.21$; $t(26) = 2.02, p = .05$), supporting Hypothesis 1. Thus, the results are consistent with our prediction that on average, individuals will create categories that are more inclusive than the categories created by interactive groups.

Discussion

It is still an open question whether these results will generalize to hierarchical categorization techniques such as affinity diagrams. An affinity diagram is similar in concept to a customer sort, but the hierarchical structure is inherent in the diagram itself rather than being inferred using cluster analysis. Affinity diagrams are particularly interesting because the process of grouping together concepts is repeated iteratively at higher and higher levels of abstraction (see Figure 1). As we mentioned previously, Hypothesis 1 has several additional implications for affinity diagrams: individuals should include a larger percentage of the requirements
available and they should be more inclusive creating categories not only at the lowest level (grouping together requirements) but also at higher levels (grouping together categories). These implications are addressed in Study 2.

**Study 2**

The goals of Study 2 were twofold. First, to see if we would observe the same patterns in group and individual categorization behavior for a hierarchical categorization technique that we observed for a single level categorization technique in Study 1. Second, to examine whether there would be enduring effects of social context for affinity diagrams due to the visual nature of the categorization task. Again, we believe these systematic differences and their enduring effects could have important implications for managers making decisions about training and the sequencing of group and individual tasks.

*Experimental Task*

The experimental task was to create a modified version of an affinity diagram. Rather than beginning with the step of defining customer requirements based on customer interview transcripts, we provided participants with customer requirements that had been previously identified. This both controlled for the set of customer requirements available to the participants and reduced the time required to complete the experimental task. For interested readers, the Appendix provides a more detailed comparison between our experimental task and similar techniques used in organizations.

Two successive experiments were run using the same design but different topics. This allowed us to demonstrate that the technique produces stable results across a variety of topics. To increase the relevance of the task for the participants and increase the generalizability of our results to an organizational context, we ran the experiments as real market research projects. We developed a situation in which participants cared about task outcomes by collaborating with two organizations who provided services used by the participants and who agreed to incorporate participant input into their decision making processes.

*Experiment 1:* In experiment 1, participants considered customer requirements for two processes within the course selection system used by MBA students: a course wait list process and a course section swap process. Participants were aware that their recommendations would be shared with the business school’s Educational Services department for use in redesigning the two processes before registration the next semester, providing the potential for student
recommendations to truly make an impact. The customer requirements used to create the diagrams during the experimental sessions were generated during eight one hour interviews with eight additional MBA students (following Burchill & Brodie, 1997). Student requirements were transcribed and evaluated to create a basic set of 24 requirements for each topic. These requirements were then pretested for understandability and rated for importance (1-3 scale) and potential to generate controversy (1-3 scale).

Experiment 2: In experiment 2, the topics were improvements to two areas of an MBA concentration (the New Product and Venture Development track): ‘Networking and Activities’ and ‘Courses and Skills.’ Participants knew that their recommendations would be shared with the track’s co-directors and the student steering committee for use in planning the track’s activities over the next year. Prior to the experimental sessions, customer requirements for the track were identified by 17 teams of two to five people who interviewed 17 members of the track. These requirements were aggregated across teams and distilled into the 24 requirements for each topic used during the experimental sessions. The final set of requirements was pretested for understandability and rated for importance (1-3 scale) and potential to generate controversy (1-3 scale).

Participants

One hundred twenty-five MBA students (88 men and 37 women; average age 29) participated in the two experiments. The 48 participants in experiment 1 participated in seven sessions ranging from four to twelve students each. The 77 participants in experiment 2 participated in eight sessions ranging from four to twenty-eight students each. Participants provided demographic information at the end of the experimental sessions. In experiment 1, participants were given lunch or snacks and a $10 gift certificate to the campus store for their participation. In experiment 2, participants were given lunch or snacks and a magazine subscription for their participation.

Procedure

The procedure was identical for experiments 1 and 2. At the beginning of each 90 minute session, the process for completing a diagram was described and demonstrated to participants using a simple example based on customer requirements for a consumer product. Following this
introduction, half of the participants created their first diagram individually and half of the participants created their first diagram as part of a four person interactive group.¹

Participants who completed their first diagram individually completed their second diagram as part of a four person group, and vice versa. Participants had the same amount of time (30 minutes) to complete each of their two diagrams. After completing each diagram, each individual wrote his or her proposed recommendation (allowing us to measure the degree of consensus within the group) and answered a few questions about the exercise (e.g., satisfaction with the proposed recommendation). Materials from the first diagram were removed before materials for the second diagram were distributed.

To demonstrate that the results are not dependent on any inherent structure in the data, the topics of the diagrams were counterbalanced, so that half of the subjects in each condition created diagrams for one topic, and half of the subjects created diagrams for a second topic. Thus, we made within subjects comparisons between group and individual diagrams across two similar cases rather than the same case. This manipulation is similar to Kim’s (1997) use of two similar survival tasks to measure the effect of task experience without exposing subjects to the same content across the two tasks.

For each of the diagram topics, participants were given a set of self-stick labels describing customer requirements, markers, and poster paper to diagram their conceptualization of student requirements for the service. Participants were told that they could use as many or as few of the labels as they wished to create their diagrams, simulating the ability of an individual to decide whether to use a piece of information or to contribute a given piece of information to a group discussion. In practice, deciding which customer requirements to use is an integral part of creating an affinity diagram (Burchill & Brodie, 1997; Shiba et al., 1993).

The first subtask in creating the affinity diagram was to review each of the customer requirements and decide whether to include it in their affinity diagram. As each requirement was selected, a Post-It note describing that requirement was affixed to a poster-sized sheet of paper. The next subtask was to combine related ideas, as in the customer sort technique. For example, two customer requirements for a vehicle, “spacious headroom” and “provides lumbar support” might be combined into a category described as “customer wants to feel comfortable in the driver’s seat.” The requirements “high quality stereo sound” and “engine runs quietly” might be

¹ With the exception of 8 control subjects in experiment 2 who completed two individual diagrams.
put into a second category with the heading “customer wants a great sound system.” The next subtask was to combine categories hierarchically. Continuing our example, the first level category “customer feels comfortable in the driver’s seat” and “customer wants a great sound system” might be combined to create the higher level category with the heading “customer wants a comfortable interior.” In turn, this second level category might be combined with other categories addressing the maneuverability of the vehicle, maintenance, or pricing.

At the experimental sessions, the affinity diagrams were described as part of the problem solving process that would enable the participants to develop a high quality recommendation for the service being addressed by the diagram. Our participants were unaware that the structure of their diagrams would be recorded or evaluated in so detailed a manner. Due to the unobtrusive nature by which our measures were collected, we believe that these process measures accurately reflect the degree to which participants were influenced by anchors, differences in their approach to the problem and the number of requirements they felt was necessary or desirable to generate a good solution.

**Design**

Three factors were examined, the social context in which the diagram was created, the sequence in which the two diagrams were created (group diagram before individual diagram, or individual diagram before group diagram), and the degree to which customer requirements were shared among individuals.

**Social Context.** Participants completed the diagram and prepared a recommendation individually in the Individual Diagram condition. In the Group Diagram condition, four participants created a single diagram together but each prepared a recommendation individually after completing the group diagram. Participants were seated at tables of four students in both conditions.

**Sequence.** Half of the participants created an individual diagram first, followed by a group diagram for a different topic. The other half of the participants created their first diagram as part of a group and their second diagram individually (again, for a different topic). In Experiment 2, as a control condition, we added a third sequence in which subjects created two individual diagrams on two different topics.

**Information Condition:** Based on research suggesting that groups tend to focus on shared information to the detriment of unshared information (Kim, 1997; Stasser, 1992; Stasser & Titus,
1985, 1987), we manipulated the degree to which customer requirements were shared among group members. To distribute the customer requirements among the members of each group, we followed the procedure for distributing information used by Stasser and Titus (1985, 1987). In the Fully Shared Information condition, each member of the group was given the same set of 24 customer requirements. In the Partially Shared Information condition, each member was given 12 shared requirements (same requirements were given to all four team members) and 3 unshared requirements (each requirement was given to only one of the four team members). Thus, in the Partially Shared Information condition, each individual was given a total of only 15 requirements, but the group as a whole had access to a total of 24 unique requirements. Across the Fully Shared and Partially Shared Information conditions, the total amount of information potentially available to the group (the number of requirements distributed to the group as a whole) was the same, but the requirements were distributed differently among individuals in the two conditions.

Both experiments utilized a fully counterbalanced 2 x 2 x 2 mixed design of (Social context: Group vs. individual) x (Sequence: Group diagram first vs. individual diagram first) x (Information condition: Fully shared vs. partially shared), with the topic of the diagram as a counterbalanced control factor. Sequence and information condition were administered between subjects and social context was a within subjects factor.

Dependent Measures

We collected dependent measures related both to the structure of the diagrams created by participants (e.g., number of available requirements used, concept differentiation, complexity and diagram structure) and to organizational considerations (e.g., consensus among group members, satisfaction, coverage of the information domain). The full set of dependent measures is described in Table 1.

Our experimental procedure resulted in the creation of a concrete artifact, a poster-sized sheet of paper, which allowed us to capture the number of requirements used, the number of requirements per category, diagram complexity, and coverage of the information domain. Dependent measures for each diagram were coded independently by two raters and inconsistencies were resolved by discussion. The free-form recommendations made by subjects after completing each diagram were content analyzed by two independent raters by assigning up

---

2 In pretests, shared and unshared requirements were rated equally in importance and controversy.
to four of ten subtopic codes based on their content.\textsuperscript{3} Again, inconsistencies were resolved by discussion. Finally, we collected ratings from each participant on measures such as their satisfaction with the recommendations they made.

In addition to collecting these survey measures and objective measures from the diagram itself, we asked the recipients of our data to rate the quality of the diagrams subjectively. Our collaborating service providers rated each diagram’s degree of problem definition structure using the Multi-Attribute Value (MAV) technique (Massey & O’Keefe, 1993), which has been used to evaluate the quality of problem definitions completed by both individuals and groups (Massey & Wallace, 1996). Structure was defined as the degree to which the diagram “reflects consideration for all relevant and unique aspects of the problem, including intrarelationships and interrelationships.” Six independent raters (two rated diagrams for experiment 1 and four rated diagrams for experiment 2) used this technique to rate diagram structure on a continuous scale ranging from 0 to 100, where lower (higher) scores indicate a lower (higher) degree of structure.

\textit{Data Analysis}

An important consideration in analyzing our dependent measures is that the individual observations for the participants in each group are likely to be correlated when individual diagrams follow group diagrams because the participants have had similar experiences. Following a method suggested by Myers, DiCecco and Lorch (1981) for comparing individual scores before and after individuals have participated in a group activity, we averaged scores for each dependent measure across the four individuals within each group unless otherwise noted (also see Stasser & Titus, 1987 for an example in a similar context). This aggregation results in two observations for each group for each of our seven dependent measures: one for the four individual diagrams and a second for the group diagram.

A second consideration is that the topics of the group and individual diagrams varied between subjects. Moreover, across our two experiments, diagrams were created for four different topics. As predicted, the topic of the diagram and the order of topics did not have significant effects on the dependent measures in either experiment. We performed GLM multivariate tests on diagram content and the order of diagram content for all 14 dependent

\textsuperscript{3} Ten subtopics were defined for each of the four diagram topics (section swap, wait list, networking and courses). For example, two subtopics addressed in wait list recommendations were ‘ease of use’ and ‘fairness/equity’ considerations.
variables (Table 1) analyzed for the group and individual diagrams. None of these variables reached marginal significance at the Bonferroni adjusted alpha level of .01 (=.10/14).\(^4\) In other words, the technique produced stable results across a variety of topics. Because the topic of the diagram had an insignificant effect on our dependent measures, we combined the data across diagram topics and experiments for the remaining analyses.

Table 1: Dependent Measures

<table>
<thead>
<tr>
<th>Dependent Measure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of Requirements Used</td>
<td>Number of requirements included in the diagram divided by the number of requirements available.</td>
</tr>
<tr>
<td>Equivalence Range</td>
<td>A measure of concept differentiation, computed as the number of requirements included in the diagram divided by the number of first level categories (average size of categories).</td>
</tr>
<tr>
<td>Diagram Complexity</td>
<td>Number of interactions identified among concepts, computed as the total number of categories defined in the diagram.</td>
</tr>
<tr>
<td>Diagram Structure</td>
<td>Qualitative rating of diagram structure by users of the data, based on the MAV technique.</td>
</tr>
<tr>
<td>Coverage of Information Domain</td>
<td>For interactive groups, number of requirements included in the diagram. For nominal groups, number of unique requirements included in any of the four individual diagrams.</td>
</tr>
<tr>
<td>Consensus in Recommendations</td>
<td>Open-ended individual recommendations were written after completing each diagram. Consensus is measured as the number of pairwise matches in subtopics (of 10 subtopics for each topic) among the four individual recommendations.</td>
</tr>
<tr>
<td>Satisfaction with Recommendation</td>
<td>Participant responses to a scale item measuring satisfaction with their recommendation.</td>
</tr>
</tbody>
</table>

For most of the analyses, we also combine information conditions when reporting the results. While we had expected to observe an interaction effect between social context and information condition, such that groups’ tendency to use fewer ideas would be exacerbated when information was not fully shared, this was not the case. There was no difference in the

\(^4\) For topic, all \(ps > .16\) except equivalence range for individual diagrams (\(p = .04\)). For order, all \(ps > .11\), with the exception of the number of ideas used in the group diagram (\(p = .07\)).
probability that shared requirements \((M = .75)\) and unshared requirements \((M = .75)\) were included in group diagrams. Moreover, with the exception of the coverage of the information domain in the individual condition, none of the dependent variables varied across information conditions at a level that reached marginal significance (Bonferroni adjusted alpha level of .01, or 10/14). Because this null result is quite interesting in light of previous research (Kim, 1997; Stasser, 1992; Stasser & Titus, 1985, 1987), and has important implications for information exchange in groups, we will discuss these findings in our discussion of the coverage of the information domain.

To examine enduring effects, we measured sequence effects across the two diagrams completed by the participants within each group. Following Laughlin and Sweeney (1977), we examine the effect of sequence using a repeated measures model, so that sequence (either individual diagram followed by group diagram or group diagram followed by individual diagram) is a main effect between subjects. This allows us to distinguish the effect of sequence from the effect of order (first diagram or second diagram) and from the effect of group/individual differences.

Results

Hypothesis 1 predicted similar effects for four dependent variables measuring diagram structure: the percentage of available requirements included in the diagram, concept differentiation, complexity, and diagram structure rating. Thus, as an overall test of Hypothesis 1, we compared individual and group results for the initial diagram created using multivariate analysis of variance (MANOVA) for these four dependent variables. The MANOVA revealed a significant effect of social context, Wilks’ lambda = .489, \(F(4,23) = 6.01, p < .005\). Supporting Hypothesis 1, these measures were consistently higher for individual diagrams than for group diagrams. In the following sections, we report the univariate results for each of these four variables.

Next, to test Hypothesis 2, we compared the same four variables across sequences using a repeated measures MANOVA with social context as a repeated factor and sequence as a between-subject factor. As predicted, the MANOVA revealed a significant main effect of sequence, Wilks’ lambda = .459, \(F(4,23) = 6.78, p < .001\). Supporting Hypothesis 2, these four

---

\(^5\) All correlations between these four variables were significant \((p<.01)\) with the exception of the correlation between complexity and equivalence range and the correlation between equivalence range and structure.
dependent measures were consistently higher when individual diagrams were created first and
group diagrams were created second than when group diagrams were created first and individual
diagrams were created second. The univariate results are described in the following sections and
are summarized in Table 2. The results of all F tests for Hypothesis 2 refer to a 2 (Social
context: Group vs. individual) x 2 (Sequence: Group-individual vs. individual-group) repeated
measures GLM in which the dependent variables for group and individual diagrams are repeated
measures.

Table 2: Enduring Effects

| Comparison of Individual Diagram First and Group Diagram First Sequences |
|-----------------------------|-------------------|-------------------|---|
| Dependent Measure           | Sequence          | First Diagram     | Second Diagram | F  |
| Percentage of Requirements Used | Individual → Group | .86 (.07)        | .85 (.11)      | 27.42** |
|                             | Group → Individual | .65 (.18)        | .62 (.09)      |   |
| Equivalence Range           | Individual → Group | 3.29 (.71)       | 3.41 (.92)     | 6.03*  |
|                             | Group → Individual | 2.97 (.75)       | 2.90 (.60)     |   |
| Diagram Complexity          | Individual → Group | 7.25 (1.87)      | 8.14 (1.96)    | 6.08*  |
|                             | Group → Individual | 7.00 (2.11)      | 6.11 (1.09)    |   |
| Diagram Structure           | Individual → Group | .20 (.47)        | .17 (.49)      | 5.27*  |
|                             | Group → Individual | -.27 (.98)       | -.18 (.55)     |   |

* significant at $p < .05$  ** significant at $p < .001$

Note: $F(1,26)$ is reported for the between subjects variable Sequence (Individual Diagram First vs.
Group Diagram First) using repeated measures GLM.

Percentage of Requirements Included. Hypothesis 1 predicted that individuals would be
more inclusive in selecting requirements for their affinity diagrams, and that individuals would
include a larger percentage of the available requirements in their diagrams than groups. The
percentage of customer requirements used was computed by counting the number of product
requirements included in each diagram and dividing the result by the number of product
requirements available. For the first diagram, individuals used a greater fraction of the requirements available ($M = .86$) than did interactive groups ($M = .65$), $F(1,26) = 15.60, p < .001$ (one-tailed), supporting Hypothesis 1. No other effects were significant.

Regarding the enduring effects of social context, Hypothesis 2 predicted that interactive groups with experience creating individual diagrams would consider more of the available requirements than interactive groups without prior experience and second, that individuals with experience creating group diagrams would consider fewer of the available requirements than individuals without prior experience. In terms of our analysis, this is a prediction that the main effect of sequence (group diagram first vs. individual diagram first) should be significant between subjects. Indeed, in our repeated measures test, we find that only the between-subject main effect for sequence was significant, $F(1, 26) = 29.41, p < .001$. The percent of available requirements in each diagram was larger when individual diagrams were created first and group diagrams were created second than when group diagrams preceded individual diagrams (Figure 2), supporting Hypothesis 2. This pattern of results suggests that while the social context shapes performance for the first diagram, enduring effects are stronger than group/individual differences in determining the number of available requirements ideas included in the second diagram.

*Figure 2: Percentage of Requirements Used in Diagrams*

Recall that the number of requirements available to individuals differed across Information conditions: in the Fully Shared Information condition, each individual was given 24 requirements, and in the Partially Shared Information condition, each individual was given 15 requirements.
**Concept Differentiation.** For the first diagram, Hypothesis 1 predicted that interactive groups would differentiate among concepts more than individuals. As in Study 1, we use Gardner's equivalence range (Gardner, 1953) to measure concept differentiation. We calculate equivalence range by dividing the number of requirements included in the diagram by the number of first level categories defined. For example, in Figure 1, the equivalence range is equal to 14/6 in Diagram A (14 ideas divided by 6 first level categories) and 14/5 in Diagram B. Therefore, in these examples, concept differentiation is greater in Diagram A (equivalence range is lower) than in Diagram B which includes more requirements per first level category. An advantage of the equivalence range measure is that it corrects for differences in the number of requirements included in each diagram.

**Figure 3: Equivalence Range**

![equivalence range graph]

Hypothesis 1 was supported only directionally. For the first diagram, equivalence range was directionally higher for individual diagrams \(M = 3.29\) than for interactive group diagrams \(M = 2.97\), \(F(1,26) = 1.38, p = .13\) (one-tailed). Hypothesis 2 predicted an enduring effect of social context for concept differentiation. We predicted, first, that interactive groups with experience creating individual diagrams would differentiate concepts less than interactive groups without prior experience, and second, that individuals with experience creating group diagrams would differentiate concepts more than individuals without prior experience. Again, Hypothesis 2 leads to the prediction that the main effect of sequence (group diagram first vs. individual diagram first) would be significant. The between-subject main effect for sequence was significant, \(F(1, 26) = 4.41, p < .05\). Overall, equivalence range was higher in both diagrams.
when individual diagrams were created first and group diagrams were created second than when group diagrams preceded individual diagrams (Figure 3). No other effects were significant.

Complexity. Our third measure of diagram structure is complexity. We operationalized diagram complexity by counting the total number of interactions defined at all levels of the diagram hierarchy. For example, in Figure 1 there are only 6 categories defined in Diagram A (all at the first level), whereas there are a total of 8 in Diagram B (5 categories at the first level, 2 at the second and 1 at the third) for the same number of requirements. As one of the implications of Hypothesis 1, we proposed that individuals would identify more interactions among requirements at all levels of hierarchy than interactive groups. For the first diagram, the number of interactions among requirements identified by individuals \( (M = 7.25) \) was not significantly greater than the number of interactions identified by interactive groups \( (M = 7.00), p > .37 \), and Hypothesis 1 was not supported for complexity.

Figure 4: Diagram Complexity

We also predicted that interactive groups with experience in creating individual diagrams would create more complex diagrams than interactive groups without prior experience, and that individuals with experience in creating group diagrams would create less complex diagrams than individuals without prior experience. Again, Hypothesis 2 predicted a main effect of sequence, which is supported by the results. Only the between-subject main effect for sequence (group diagram first vs. individual diagram first) was significant, \( F(1, 26) = 5.91, p < .05 \). A larger total number of interactions among requirements were identified in each diagram when individual
diagrams were created first and group diagrams were created second than when group diagrams preceded individual diagrams (Figure 4).

**Diagram Structure.** As we have defined it, diagram structure reflects whether interactions among requirements are perceived to have been determined using clear rules. We asked the users of our data, the service providers we worked with in these two service improvement studies, to rate the degree of structure for each diagram. We predicted that because groups differentiate concepts more and identify fewer interactions among categories, group diagrams should appear to have less inherent structure than individual diagrams. This prediction was marginally supported. For the first diagram, the average (normalized) structure ratings for individual diagrams \( M = .20 \) were marginally higher than the ratings for interactive group diagrams \( M = -.27 \), \( F(1,26) = 2.62, p = .06 \) (one-tailed).

**Figure 5: Diagram Structure**

![Diagram Structure](image)

Again, Hypothesis 2 suggested that the sequence effect (group diagram first vs. individual diagram first) for diagram structure should be significant. As predicted, only the between-subject main effect for sequence was significant, \( F(1, 26) = 5.16, p < .05 \). Diagram structure was given higher ratings by users of the data for both diagrams when individual diagrams preceded group diagrams than when group diagrams preceded individual diagrams (Figure 5).

---

7 The effective reliability of our raters using the Spearman-Brown formula ranged from 62% to 77% and averaged 70% across the four diagram topics. This level of reliability compares favorably with correlations documented by Massey and Wallace (1996).
Analysis of these four structural dimensions shows a consistent pattern of results, providing support for Hypothesis 1 and Hypothesis 2. However, differences in the structure of the diagrams themselves are only one of the considerations for managers when making decisions about product development processes. There are other considerations in addition to the outcomes of the task that differentiate individual and group work (Sutton & Hargadon, 1996). In the next section, we discuss additional organizational considerations relevant to the use of affinity diagrams.

Organizational Considerations

When deciding whether tasks should be performed by individuals or groups, it is important to consider factors such as how the level of consensus among team members varies after completing diagrams individually or in a group, the participants’ satisfaction with the task and output, and the comprehensiveness of the analysis (coverage of the information domain). In this section, we review the results of Study 2 with respect to these three dimensions. These results are summarized in Table 3.

Table 3: Overall Group vs. Individual Differences

<table>
<thead>
<tr>
<th>Overall Comparison of Diagrams Created by Groups and Individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Measure</td>
</tr>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td>Consensus</td>
</tr>
<tr>
<td>Satisfaction</td>
</tr>
<tr>
<td>Coverage of Information Domain</td>
</tr>
</tbody>
</table>

* significant at $p < .05$  ** significant at $p < .001$

Note: For consensus and satisfaction, $F(1,26)$ is reported for the between subjects variable Sequence (Individual Diagram First vs. Group Diagram First) using repeated measures GLM. For coverage, $F(1,24)$ is reported for the between subjects variables Sequence and Information (Fully Shared vs. Partially Shared) using repeated measures GLM.

Consensus. One reason for using teams is that organizational buy-in to the problem formulation after it has been completed is often crucial. Group performance of a task may lead to higher consensus, coordination of activities, or better mutual understanding of a problem among group members via the development of shared mental models (Klimoski & Mohammed, 1994; Larson & Christensen, 1993).
One way to quantify the degree of consensus among group members is to compare the recommendations of individual group members made following the creation of group affinity diagrams and following the creation of individual affinity diagrams. While we did not explicitly ask participants to arrive at a consensus as part of the experimental task, we asked subjects to respond to the open ended question “What do you think are the most significant student requirements for (diagram topic)?” after completing each diagram. We content analyzed the recommendations (see Dependent Measures) and analyzed the number of pairwise matches in subtopics across subjects within each nominal group or interactive group.

As predicted, our measure of consensus was systematically higher among recommendations written after participants had created a group diagram ($M = 6.78$) than after the same four participants had created diagrams individually ($M = 4.89$; see Table 3). The within-subject main effect for social context was significant, $F(1, 25) = 4.89, p < .05$, and no other effects, including the between-subjects main effect for sequence, were significant, $p > .25$. This suggests that creating a single diagram as a group may have some benefits for creating buy-in to a common solution relative to investing the same amount of time creating four individual diagrams.

**Satisfaction.** In addition to task output, group member satisfaction is an important criterion for team effectiveness (Hackman, 1987) and is a valued result in organizational settings (Sutton & Hargadon, 1996). For both sequences (group diagram first or individual diagram first), satisfaction with final recommendations was significantly higher after completing individual diagrams ($M = 5.73$) than after completing group diagrams ($M = 5.19$), $F(1, 24) = 4.20, p < .05$, with no significant interactions for other variables. Satisfaction with recommendations was significantly correlated with perceived sufficiency of information ($r = .59, p < .001$ following individual diagrams; $r = .73, p < .001$ following group diagrams). After completing group diagrams, satisfaction with recommendations was significantly correlated with an index of four items measuring satisfaction with the group experience ($r = .63, p < .001$).

---

8 Perceived sufficiency of information was measured by response to the item “I feel that I have enough information to make this recommendation.” Though half of those completing individual diagrams (those in the partially shared information condition) had fewer ideas available to them than to interactive groups, perceived sufficiency of information was higher for individual diagrams ($M = 5.37$) than for group diagrams ($M = 4.82$), $F(1, 24) = 5.25, p < .05$. Interestingly, this measure did not differ by information condition for either individuals or groups (though a null effect for information condition is counterintuitive, it is not surprising; also see Slovic, Fischoff & Lichtenstein, 1982).
Coverage of the Information Domain. Because each individual adds the benefit of their own knowledge and experiences to the group, the potential coverage of the information domain should be higher for a group than for the same number of individuals. However, the process of combining individual inputs is not a simple additive one. While potential coverage of the information domain may be higher in groups, actual coverage tends to be lower than potential coverage in interactive group settings (Walsh et al., 1988). One reason this may happen is because groups tend to focus more on information shared by all group members than on information that is known by only one or a few group members (Kim, 1997; Stasser & Titus, 1985, 1987; Stasser 1992). Thus, coverage of the information domain is an especially important consideration when individual group members have access to specialized information, such as in cross-functional teams. Given that much of the product development literature asserts that best-practice firms use cross-functional teams (Griffin, 1997), the issue is particularly relevant for product development.

We compared the aggregate coverage of the four individual diagrams created by each group with the coverage of the group diagram they created for a different topic. We calculated the coverage of the information domain for individuals by counting the number of unique requirements included by any of the four members of each group in their individual diagrams. Coverage for group diagrams is simply the number of requirements included in the group diagram. Coverage was systematically higher for the four individual diagrams ($M = 21.82$) than for the interactive group diagrams ($M = 18.00$), $F(1,24) = 30.51, p < .001$, indicating that the diversity of requirements considered when four group members create diagrams independently is significantly higher than when the same four people create a single diagram together.

We also predicted that the lower coverage of the information domain realized by interactive groups relative to the same number of individuals would be exacerbated when the information given to group members was only partially shared. This hypothesis was not supported. The main effect of social context was indeed qualified by a significant two-way interaction between social context and information condition, $F(1, 24) = 8.57, p < .01$. However, this was because individual coverage was significantly higher in the Fully Shared Information condition ($M = 23.21$) than in the Partially Shared Information condition ($M = 20.43, t(26) = 3.97, p < .001$) for both sequences. This effect can be explained by the number of ideas available to each individual. In the Fully Shared Information condition, each of the 24 ideas was
available to all four individuals, whereas each of the 12 unshared ideas was only available to one of the four members in the Partially Shared condition. Thus, the probability that each requirement would be included in one of the four group members’ diagrams was higher when the requirement was available to all four group members than when it was only available to one individual (Stasser & Titus, 1985, 1987). No other interactions were significant.\(^9\)

This null effect is actually quite significant in practical terms, because it suggests that making ideas explicitly available (e.g., the written requirements we gave to our subjects) prevents groups from focusing on shared information to the detriment of unshared information. In previous tests (e.g., Stasser & Titus, 1985; 1987), information was read by each subject prior to group discussion but no written documents were available to the subjects during the discussion. Our results suggest that groups’ tendency to focus on shared information might be prevented if, for example, each subject put the ideas they wanted to contribute to a group discussion into a format easily shared with others (e.g., putting customer requirements on Post-It notes, as is normally done when using the affinity diagram technique).

**Discussion**

From the affinity diagrams created by participants, we collected a varied set of dependent measures to capture their mental representations of the relationships between customer requirements. In addition to counting the number of requirements included in each diagram, we computed measures of concept differentiation, diagram complexity, and diagram structure. These measures focused not on the content of the diagrams, but on the way relationships between concepts were defined by groups and individuals. Across the two experiments and a total of four topics, our results consistently show that the use of the affinity diagram technique is both social context dependent and history dependent.

In some cases, we observed systematic group/individual differences regardless of past experience with the affinity diagram technique. For example, participants’ satisfaction with their recommendations was consistently higher after completing individual diagrams, but group member buy-in to a common solution was higher after the four participants had created a single group diagram than after they had created four individual diagrams. In other cases, we observed

---

\(^9\) While the between subjects main effect for information condition was not significant \((p > .16)\), the between-subjects main effect for sequence was significant, \(F(1, 24) = 17.28, p < .001\). This effect can be explained by the larger number of ideas considered in both diagrams when the individual diagram was created first. However, for each sequence, the main effect of social context was significant in a paired t-test \((p < .01)\).
strong sequence effects. Participants who completed individual diagrams first included a higher percentage of the available customer needs in both of their diagrams than participants who completed group diagrams first. Across both first and second diagrams, those who created individual diagrams first categorized customer needs more inclusively (lower concept differentiation), defined a larger number of interactions among customer needs (more complexity), and created diagrams rated higher in structure by subject area experts.

How can we be sure that it was experience with the technique in a group or individual setting that caused these enduring effects? By design, we have ruled out order (first or second diagram) as the cause of the sequence effects. If the effects were due to fatigue or practice with the experimental task, then the first and second diagrams created by both individuals and groups should differ systematically. However, we see very different trends for group and individual diagrams produced second relative to diagrams produced first.

Another finding that argues against an order effect is the high consistency between the two individual diagrams produced in the Individual-Individual control condition. For example, there was no difference in the percentage of available requirements included by control subjects in their first ($M = .87$) and second diagrams ($M = .89$), $p > .53$. Two interesting comparisons can be made between individual diagrams in the test conditions and individual diagrams in the control condition. Specifically, the percentage of available requirements included by the control subjects in their first diagram ($M = .85$) did not differ from the percentage of available requirements included in the first diagram by individual in the test condition ($M = .83$), $t(14) = .43, p > .67$. In contrast, the percentage of available requirements control subjects included in their second diagrams ($M = .89$) was significantly greater than the percentage of available requirements individuals in the test condition included in their second diagrams, which they completed after having worked on a group diagram ($M = .63$), $t(14) = 4.21, p < .001$. Other measures (equivalence range, complexity and structure) show the same pattern of insignificant differences across the two individual diagrams in the control condition and between first diagrams created by individuals in the test condition and the control condition. These results suggest that the structure of affinity diagrams is history dependent rather than order dependent.

Why should the experience of having created an affinity diagram individually or as part of a group make so much difference in the way the technique is used subsequently? Previous research examining group to individual and individual to group transfers between tasks has
found mixed results (Laughlin & Sweeney, 1977). We suggest that for the affinity diagramming technique, in contrast to a less visual task like problem solving, strong cues are provided by the visual output (the affinity diagram) that is created. Even if these cues are not physically present, subjects’ memory for visual characteristics of the first diagram provides a mental model that shapes their performance expectations for the second diagram. Thus, we suggest that the visual characteristics of the diagram (e.g., number of requirements included, size of the categories created, complexity) should be most susceptible to carry-over effects across diagrams. An interesting implication, which will be tested in Study 3, is that merely observing an example of a categorization technique applied to one topic may influence the structure of the categories perceived for another topic. In other words, the category structure does not emerge from the concepts themselves, but from the subjects’ mental model of the categorization structure.

**Hypothesis 3:** The visual structure of a categorization technique applied to one topic will shape the way people categorize ideas for an unrelated topic.

**Study 3**

The goal of Study 3 was to test the mechanism for the enduring effects of social context observed in Study 2. Hypothesis 3 suggests that seeing an example of a categorization technique applied to one topic will shape the way people categorize ideas for a completely different topic. To test this hypothesis, subjects were given the same instructions for creating an affinity diagram, but two different examples of affinity diagrams. We expected the characteristics of the diagram model subjects were given to alter their mental model of an affinity diagram and change the goals subjects had for creating their own affinity diagrams. Because we thought this effect might be moderated by familiarity with the technique, we measured subjects’ prior experience using affinity diagrams and similar techniques.

As in Study 2, we examined dependent measures reflecting the structure of the diagrams subjects created, such as the number of requirements included in the diagram, the size of the categories created and the number of interactions defined among requirements.

**Subjects**

Thirty-six MBA students participated in the study, and were paid $5 for their participation. Nineteen subjects were randomly assigned to the low structure condition and seventeen subjects were randomly assigned to the high structure condition. Eleven of the
subjects (six of nineteen subjects in the low structure condition and five of seventeen in the high structure condition) had previous experience using affinity diagrams or similar techniques.

Procedure

Subjects were given one of two model diagrams analyzing customer requirements for a consumer product (shaving cream) as an example. The two model diagrams used as stimuli are shown in Figure 6. The models differed in the percentage of the available requirements used in the diagram, the number of requirements per category, and in their overall complexity (number of interactions defined among requirements). The low structure model was designed to exemplify the characteristics of the first diagrams created by groups in Study 2, using fewer of the available requirements, including fewer requirements in each category, and identifying fewer interactions among requirements. The high structure model was designed to exemplify the characteristics of the first diagrams created by individuals in Study 2, using a larger percentage of the available requirements, including more requirements in each category, and identifying more interactions among requirements fewer ideas per category.

All subjects created their affinity diagrams individually. Subjects in both conditions were given identical written instructions for completing their diagrams, the same instructions that had been used in Study 2. To create their own affinity diagrams on a different topic, subjects were given 20 customer requirements for MBA Tracks (concentrations). These requirements had been identified in previously conducted student interviews, and were a subset of the requirements that had been used in Study 2. The experimental task took about 30 minutes to complete.

Results

Results were analyzed using a GLM model with one factor identifying the diagram model given to the subject (low structure model or high structure model) and using the subject’s prior experience with affinity diagrams (prior experience or no prior experience) as a covariate. Subjects’ previous experience with similar techniques did not differ across conditions ($p > .50$).
Figure 6: Diagram Models

Low Structure Model:

Example of an Affinity Diagram Created by a Shaving Cream Customer

Previously identified customer requirements:

- Minimizes burning sensation
- Minimizes cuts
- Minimizes stubble
- Bottle doesn’t leave rust ring
- Doesn’t dry out skin

Packaging is efficient
Maximizes uses per bottle
Makes skin feel smooth
Avoid making a mess
Bottle doesn’t leave rust ring
Packaging prevents accidental discharge

Makes skin feel good after shaving
Minimizes stubble
Doesn’t dry out skin

Makes shaving a pleasant experience
Minimizes pain
Minimizes burning sensation
Minimizes cuts
Smells clean

The customer has arranged and labeled key requirements for the product to create an affinity diagram.

High Structure Model:

Example of an Affinity Diagram Created by a Shaving Cream Customer

Previously identified customer requirements:

- Maximizes uses per bottle
- Minimizes cuts
- Minimizes stubble
- Bottle doesn’t leave rust ring
- Doesn’t dry out skin

Packaging is convenient
Maximizes uses per bottle
Dispenser minimizes waste
Product is easy-to-use and effective
Bottle doesn’t leave rust ring
Packaging prevents accidental discharge

Provides components of a good shave
Makes skin feel good after shaving
Minimizes stubble
Doesn’t dry out skin
Minimizes burning sensation

Makes shaving a pleasant experience
Minimizes pain
Smells clean
Minimizes cuts

The customer has arranged and labeled key requirements for the product to create an affinity diagram.
As predicted, the number of requirements used by subjects varied significantly based on the model they had been given. Subjects who were given the low structure model including fewer of the available requirements created diagrams using only 75% of the requirements ($M = 14.95$ of 20 ideas), while subjects who saw the high structure model including more of the available requirements created diagrams using 86% of the requirements ($M = 17.29$ of 20 ideas; $F(1,33) = 4.23, p < .05$). This difference occurred despite the fact that subjects in both conditions were given identical written and verbal instructions.

Equivalence range, the average number of requirements included in each category, also differed marginally based on the model subjects had been given. Subjects who saw the low structure model using smaller categories created diagrams with an average of 3.03 requirements per category, while subjects who saw the high structure model using larger categories created diagrams with an average of 3.95 requirements per category ($F(1,33) = 3.25, p < .10$). However, the complexity of the diagram (number of interactions defined among the requirements) did not differ significantly depending on the model received ($M = 6.63$ vs. $M = 6.82, p > .75$).

The covariate, previous experience, was not significant for either the number of requirements used nor the number of requirements per category ($ps > .41$). However, diagram complexity was significantly predicted by previous experience. Those who had previous experience with the technique created diagrams with significantly more interactions among requirements ($M = 7.82$) than those who did not have previous experience ($M = 6.24; F(1,33) = 4.06, p < .05$). One potential explanation for this result is that both of the models had a rather low level of complexity. Perhaps those with previous experience had mental models of diagrams that were relatively complex.

Discussion

These results provide additional evidence that one’s expectations about the outcome (based on an example of the technique for a different topic, for example) influence the category structure perceived among concepts. This reinforces our earlier finding that the underlying similarities perceived among the concepts being categorized do not fully determine the categorization structure created. Categorization behavior is malleable, potentially influenced not only by the social context in which a categorization technique is used, but also by expectations about the finished product.
One important implication of these findings is that the structural characteristics of examples given to people while they are learning to use such a technique are critical. Starting off with a simple example could lead to simple results when the technique is used in practice. Moreover, once people have been influenced by the structural characteristics of an example, their own diagrams may reinforce these same characteristics.

**Conclusion**

Our experiments captured process artifacts that enabled us to compare the manner in which groups and individuals categorized the same set of customer needs. The findings of the three studies reported here suggest that categorization behavior is quite sensitive to contextual factors. Our findings support both group/individual differences in categorization behavior and enduring effects of prior experience with a categorization technique. Moreover, our results reinforce the importance of visual task instructions as models for performance.

We observed systematic effects of social context for the customer sort in Study 1 and for the first affinity diagram created in Study 2. In both cases, few performance cues were available, and social context had a strong effect. In contrast, for the second affinity diagram created in Study 2 and for the affinity diagram created in Study 3, individuals had performance cues from their previous experience with the technique or from the example they had been given. This suggests that under certain conditions, other factors may be more important than social context. We have argued that an individual’s initial prior experience with the same (or a similar) task is a mechanism for forming a mental representation and for setting expectations and performance standards. The pattern of results we found in Studies 1, 2 and 3 suggests that social context is particularly important when a new task is learned, and that the most visual aspects of task performance may be particularly susceptible to carryover effects. Adding confidence to this explanation, Study 3 demonstrates that visual characteristics of affinity diagrams can be shaped by exposure to examples which vary along specific dimensions.

Our finding that individuals and groups categorize differently suggests that social context deserves managerial consideration. Whether people are trained to use a categorization technique individually or in a small group setting can have lasting effects on the way they use the technique. Even when they were given exactly the same instructions, people who learned the technique individually formed different habits from those who learned the technique in a group setting. To the extent that there are process losses or gains inherent in certain group or individual
tasks, our results suggest that they may affect performance on future (similar) tasks, even in a
different social context.

We suggest that systematic differences in the learning and use of categorization
techniques by groups and individuals can be leveraged when determining an overall training
strategy or planning the product development process. If the social context in which a technique
is learned has lasting effects on the use of the technique, firms might consider training teams
individually to retain individual-like traits in team settings, or vice versa. Alternatively, a single
task might be divided into phases of individual and group activity. For example, inputs to group
workshops might be produced during individual pre-workshop activities, or group members
might use outputs produced during the workshops for individual follow-up activities. Based on
the systematic differences predicted between individual group members and groups, managers
can design a sequence of activities to achieve appropriate organizational goals.

Of course there are no simple answers, and many objectives must be balanced. The
process measures we captured are relevant to several potential organizational goals: diversity of
ideas, consensus, degree of problem structure, and satisfaction of the participants. Because the
relative importance of these criteria is likely to vary based on the problem solving objectives
(e.g., the phase of a project), we do not suggest a particular strategy such as individual learning
followed by group performance, or vice versa. Rather, we suggest that social context deserves
managerial attention.
References


among tree experts: Do all roads lead to Rome? *Cognitive Psychology*, 32(1), 49-96.


Appendix

A Note On The Technique Used

While our experimental procedures reflected the process of creating affinity diagrams prescribed by the Center for Quality of Management (the LP Method\textsuperscript{10}) as closely as possible, we made two modifications. First, the LP Method is not designed to be done repeatedly with the goal of creating easily analyzed output. The entire process of creating an LP diagram, including interpretation of customer requirements, may take three to four hours. Moreover, even if groups are given identical transcripts of customer interviews, the requirements used in the diagrams are likely to vary widely across diagrams. Thus, to increase the comparability of the diagrams for this experiment, customer requirements were defined prior to the affinity diagram sessions and distributed to the participants on self-stick labels for use in creating their diagrams. As in the LP Method, participants using our modified technique were not required to use all of the available requirements.

The second difference between our experimental task and the LP Method is that the LP Method limits the size of categories to no more than three requirements per category, minimizing the structural differences between diagrams. Using our modified technique, participants were allowed to include as many requirements as they felt were appropriate in each category. However, as in traditional affinity diagrams, participants grouped together customer requirements at increasing levels of abstraction using a poster-sized sheet of paper and markers to create a visual representation of their category structure.

\textsuperscript{10} The Language Processing Method (LP; see Burchill & Brodie, 1997) and the KJ Method (developed by Jiro Kawakita; see Shiba et al., 1993) are functionally equivalent methods for creating affinity diagrams.