

1997

# Context and order effects in self-administered surveys

Susan R. Rosenberg  
*San Jose State University*

Follow this and additional works at: [https://scholarworks.sjsu.edu/etd\\_theses](https://scholarworks.sjsu.edu/etd_theses)

---

## Recommended Citation

Rosenberg, Susan R., "Context and order effects in self-administered surveys" (1997). *Master's Theses*. 1527.  
DOI: <https://doi.org/10.31979/etd.4n6n-vxdz>  
[https://scholarworks.sjsu.edu/etd\\_theses/1527](https://scholarworks.sjsu.edu/etd_theses/1527)

This Thesis is brought to you for free and open access by the Master's Theses and Graduate Research at SJSU ScholarWorks. It has been accepted for inclusion in Master's Theses by an authorized administrator of SJSU ScholarWorks. For more information, please contact [scholarworks@sjsu.edu](mailto:scholarworks@sjsu.edu).

## INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

**The quality of this reproduction is dependent upon the quality of the copy submitted.** Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

# UMI

A Bell & Howell Information Company  
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA  
313/761-4700 800/521-0600



**CONTEXT AND ORDER EFFECTS IN SELF-ADMINISTERED SURVEYS**

**A Thesis**

**Presented to**

**the Faculty of the Department of Psychology**

**San Jose State University**

**In Partial Fulfillment**

**of the Requirements for the Degree**

**Master of Science**

**by**

**Susan R. Rosenberg**

**August, 1997**

**UMI Number: 1386228**

---

**UMI Microform 1386228**  
**Copyright 1997, by UMI Company. All rights reserved.**

**This microform edition is protected against unauthorized  
copying under Title 17, United States Code.**

---

**UMI**  
**300 North Zeeb Road**  
**Ann Arbor, MI 48103**

© 1997

Susan R. Rosenberg

**ALL RIGHTS RESERVED**


APPROVED FOR THE DEPARTMENT OF PSYCHOLOGY



Dr. Howard Tokunaga

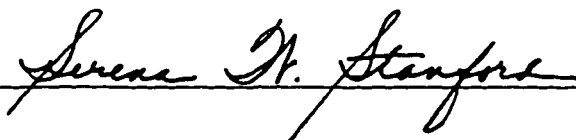


Dr. Rebecca Malatesta



Dr. Susan Bachman, Kaiser Permanente

APPROVED FOR THE UNIVERSITY



## ABSTRACT

### CONTEXT AND ORDER EFFECTS IN SELF-ADMINISTERED SURVEYS

by Susan R. Rosenberg

The purpose of this study was to examine the effect of context on self-report responses. Eight different surveys, each consisting of a set of general membership items and a module of items about a specific health care experience, were administered to a sample of health plan members. The surveys consisted of four different service types: laboratory, pharmacy, radiology, and outpatient and two order types, general items first or specific items first. Across all survey types, 4,792 respondents completed a survey.

Results of ANOVAs showed that respondents rated general, ambiguous items differently when different items were placed around them, but specific items were not affected by changing context. Correlations showed that the relationship between general and specific items is stronger when specific items are presented first. These findings have implications for organizations using surveys to make policy decisions.



## TABLE OF CONTENTS

SECTION	PAGE
INTRODUCTION .....	3
Context and Order Effects .....	5
Cognitive Processes Underlying Attitude Formation and Retrieval .....	6
Comprehension .....	7
Retrieval of Relevant Attitudes .....	7
Forming a Judgment .....	8
Selecting a Response .....	9
Factors That Influence the Occurrence of Context Effects .....	9
Item Content --General vs Specific-Concrete vs Abstract .....	10
Relationships Among Items .....	12
Part-Whole item relationships .....	13
Part-Part item relationships .....	16
Psychological Influences--Attitude Intensity & Issue Involvement..	17
Respondent Characteristics .....	20
Mode of Administration .....	22
Summary of the Literature .....	24
Background Information on the Current Study .....	26
Purpose of This Study .....	27

<b>METHOD</b> .....	29
<b>Participants</b> .....	29
<b>Materials</b> .....	30
<b>Dependent Variables</b> .....	33
<b>Context Variables</b> .....	34
<b>Procedure</b> .....	38
<b>RESULTS</b> .....	38
<b>DISCUSSION</b> .....	52
<b>Implications of This Study</b> .....	60
<b>Strengths and Limitations</b> .....	61
<b>Future Research</b> .....	62
<b>REFERENCES</b> .....	64

## LIST OF TABLES

TABLE	PAGE
1. Survey Outgo, Returns and Response Rates by Service Type .....	31
2. Number of Returns by Order Within Service Type .....	32
3. Descriptive Statistics For General Dependent Measures .....	42
4. Analysis of Variance for Global Service Measure .....	44
5. Analysis of Variance for Global Access Measure .....	45
6. Analysis of Variance for Overall Experience Measure .....	46
7. Descriptive Statistics For Specific Dependent Measures .....	49
8. One-Way Analysis of Variance for Service Specific Modules .....	50
9. Correlations of Overall Visit Satisfaction With General Dependent Measures ....	51

Running head: CONTEXT AND ORDER EFFECTS

Context and Order Effects in Self-Administered Surveys

Susan R. Rosenberg

San Jose State University

### Abstract

The purpose of this study was to examine the effect of context on self-report responses. Eight different surveys, each consisting of a set of general membership items and a module of items about a specific health care experience, were administered to a sample of health plan members. The surveys consisted of four different service types: laboratory, pharmacy, radiology, and outpatient and two order types, general items first or specific items first. Across all survey types, 4,792 respondents completed a survey. Results of ANOVAs showed that respondents rated general, ambiguous items differently when different items were placed around them, but specific items were not affected by changing context. Correlations showed that the relationship between general and specific items is stronger when specific items are presented first. These findings have implications for organizations using surveys to make policy decisions.

### Context and Order Effects in Self-Administered Surveys

If you want to know what someone thinks or feels about something, just ask them. This bit of advice seems, on the surface, to be the obvious solution. But, how easy is it really to obtain honest, thoughtful answers about attitudes, beliefs, or opinions; and what is the best way to ask in order to solicit honesty and candor from respondents? These are questions that have been plaguing and intriguing survey researchers for years, yet there seem to be no easy answers. Self-report measures in their variety of forms and modes have become popular methods to solicit information from people about a variety of issues; yet, despite their popularity, there has been little applied research on the complex interaction of factors that influence self-report responses (Harrison & McLaughlin, 1993). The purpose of this study is to examine the effect of context on self-report responses.

Research has shown that there are a host of “response effects” that can obscure or skew respondents’ true attitudes, even without their conscious knowledge that these processes are occurring (Krosnick & Schuman, 1988). Response effects can be variously defined as response consistencies that are independent of the construct under investigation (Harrison & McLaughlin, 1993) or, more precisely, systematic variations in self-reporting of attitudes due to properties of the questionnaire itself such as question content, order, context, response options, or mode of survey administration (Krosnick & Schuman, 1988). More simply stated, varying the properties of the questionnaire itself such as subtle changes in the wording of questions; changes in the context of the items, either by

changing the order of items or by including completely different sets of items with a static set of target questions; differing response options, for example, using a poor to excellent scale rather than a satisfaction or agreement scale; or varying survey administration procedures such as using telephone interviews instead of self-administered mail-back surveys can all introduce systematic bias in responses that is unrelated to the construct that researchers believe they are measuring.

Research has, for the most part, failed to produce well developed models for detecting, measuring, and predicting the occurrence of response effects and, in the case of context effects, has frequently produced conflicting findings (Krosnick & Alwin, 1987). For example, some researchers have found evidence that supports their hypotheses but are then unable to replicate these findings. Other studies have obtained results in the opposite direction from what was expected. Details on this research will be presented in subsequent sections. Despite these frustrating results, the fact remains that it is of major importance for survey designers to understand the properties of their instrument and the underlying processes influencing respondent question answering behavior since the very nature and complexity of response effects makes them difficult to predict and to replicate; hence, they may pose a serious threat to construct validity. That is, the instrument may not be measuring the desired construct as much as it is measuring the effects of contextual cues. This problem has serious implications since important policy decisions are frequently made on the basis of survey responses (Harrison & McLaughlin, 1993;

Schuman & Presser, 1981). It is clear that there are no simplistic explanations that can aid in the prediction or prevention of response effects; rather, it appears that psychological factors and respondent characteristics may be interacting with properties of the survey itself to exert influence on survey response processes (Tourangeau, 1987). A particularly complex and pervasive type of response effect are context effects. These will be the focus of this study.

#### Context and Order Effects

Context effects can be generally defined as any interpretation that a subject places on an item as a result of its relation to the other elements of the survey. A growing body of research supports the notion that systematic variations in responses on self-report instruments do, in fact, frequently occur when item context is altered (Harrison & McLaughlin, 1993). For example, when respondents are asked to answer two questions of equal specificity that require them to render a judgment of fairness or equity, having answered the first question in a certain way, respondents will often answer a second related question in the same direction. On the other hand, if respondents are asked to answer two seemingly related questions, the first of which is specific and the other general, respondents will frequently eliminate the specific example from their judgment process and answer the general question in a way that seems incongruous with their earlier answer (Schuman & Presser, 1981). Order effects are one type of context effect in which earlier items influence responses to later items. It is important to note, however, that



survey respondents may also be influenced by other aspects of the survey context such as introductions, instructions, and stems (Tourangeau & Rasinski, 1988). Additionally, question content, relationships among items, and mode of survey administration may play a role in determining whether context effects will occur and how they will manifest. Likewise, psychological factors such as social norms, attitude intensity, and issue involvement, as well as respondent characteristics such as education level can influence the occurrence of context effects (Harrison & McLaughlin, 1993; Schuman & Presser, 1981; Tourangeau & Rasinski, 1988). Subsequent sections of this paper will provide detailed explanations and examples of research on each of these influences.

#### Cognitive Processes Underlying Attitude Formation and Retrieval

Context effects and reasons for their occurrence can best be understood by examining the underlying cognitive processes involved in answering attitude questions. Cognitive psychologists assert that attitude formation and memory are organized around relatively stable cognitive structures referred to as scripts or schemas (Tourangeau, 1987; Tourangeau & Rasinski, 1988). These scripts, which are analogous to attitudes, can be thought of as “networks of interrelated beliefs”. Beliefs, in turn, result from a combination of specific experiences, feelings, values, and knowledge about things or common situations such as going to the doctor. How these relevant attitudes are accessed and interpreted in light of the question being asked will determine how the question will be answered. Tourangeau (1987) posits that there are four major stages a respondent goes

through when answering an attitude question. Each stage provides opportunities for context effects to occur. Understanding the cognitive processes at each stage that are thought to lead to context effects may enhance their predictability.

### Comprehension

In the comprehension stage, the challenge is to understand the question. Sometimes it will be obvious to the respondent what attitude or belief the question is trying to measure, and sometimes it will not be so obvious. When the topic is familiar and the question is unambiguous, the respondent can easily access the relevant attitude and answering the question becomes virtually automatic. However, when the topic is unfamiliar or the question content is too general or unclear, then respondents must search for cues to help them interpret what is being asked. In this case, context, which may be set by earlier questions, may be used as an “interpretive framework” to determine the meaning of subsequent questions (Tourangeau, 1987; Tourangeau & Rasinski, 1988).

### Retrieval of Relevant Attitudes

In the retrieval stage, the respondent calls from memory attitudes, beliefs, and feelings about the issues under investigation. Attitude accessibility plays an important role in determining which memories will be retrieved. Variables that are likely to influence attitude accessibility are expertise, issue involvement, and recent retrieval of the relevant attitude. Expertise refers to knowledge about the topic; issue involvement refers to caring about the topic. Recent retrieval stems from recent experiences with or thoughts about

the topic (Tourangeau, 1987; Tourangeau & Rasinski, 1988). When asked about a certain topic, respondents are unlikely to search for all relevant information in their memory banks; rather, they are likely to think about the topic only as long as is necessary for something to come to mind that is relevant to answering the question. At this stage, context, set by earlier questions, may make certain attitudes more accessible (Bishop, 1987). "Because respondents are unlikely to retrieve all their beliefs about an issue, the retrieval stage can be seen as a kind of sampling process that over-represents the most accessible beliefs or situational cues" (Tourangeau & Rasinski, 1988, p. 300).

### Forming a Judgment

In the judgment stage, the respondent examines the recalled information in light of his or her understanding of the question and renders a judgment. Different types of attitude questions present respondents with varying degrees of difficulty in rendering a judgment. In some cases, the question may be asking the respondent to simply agree or disagree with a statement that triggers a highly accessible existing attitude. These types of judgments most likely will not be affected by context effects. At the other extreme, some attitude questions require respondents to make complex judgments like fairness or judgments on multidimensional issues (Tourangeau, 1987; Tourangeau & Rasinski, 1988). "In such cases respondents may simplify their task by focusing on dimensions and standards of comparison that readily come to mind, and context can determine which dimensions or standards come to mind" (Tourangeau & Rasinski, 1988, p. 310).

### Selecting a Response

At this final stage, the respondent must select the most appropriate response from the choices that are given. This stage frequently produces response effects, not only because respondents are usually limited to a finite set of choices in which to express their opinions, but they must do so in a way that makes them feel comfortable. This may lead respondents to edit their judgments in order to answer in a way that they consider to be socially desirable. Additionally, context may once again play a role since respondents frequently will attempt to respond in ways that are consistent with their earlier responses (Tourangeau, 1987; Tourangeau & Rasinski, 1988).

### Factors That Influence the Occurrence of Context Effects

Introductory sections of this paper have alluded to a plethora of factors that can influence the occurrence of context effects. In order to understand which of these factors may be contributing undesirable variation to one's own research study, it is particularly important to examine each potential influence and understand which elements of the survey or survey items may elicit these influences. Subsequent sections will provide a comprehensive look at the research on these various factors.

When context effects occur, they can manifest as either consistency effects or contrast effects. Consistency effects, also known as assimilation effects or carry-over effects, induce respondents to answer subsequent questions in a way that is consistent with earlier responses. Contrast effects, also known as backfire effects, manifest when

respondents answer a target question in a manner which seems incongruous with earlier related context questions (Schuman & Presser, 1981; Teas & Wong, 1992). Research aimed at discovering the impetus for these context effects has produced conflicting results (Schuman & Presser, 1981). As noted earlier, a number of variables have been identified as contributors to the propensity for their occurrence. Certain social norms (to be discussed later) come into play here interacting with the perceived relationships among items to exert influence on the respondent's propensity to answer in a manner which is either consistent or inconsistent with earlier responses (Schuman & Presser, 1981).

#### Item Content--General Versus Specific-Concrete Versus Abstract

This section will examine how characteristics of item content can influence respondents to use context set by earlier questions to answer subsequent questions. Research has shown that general items, for example, items that are asking for an overall rating of some construct such as service, are far more vulnerable to context effects than items that are asking about a specific aspect of the construct. In fact, in research studies where general and specific items are reversed, it is always the responses to general items that change due to order; specific items are unaffected (Schuman & Presser, 1981). Benton and Daly (1993) maintain that it is not advisable for survey researchers to ask a general question first that requires respondents to consider a multitude of factors. Instead, it is preferable to ask groups of questions that deal with specific aspects of the construct being measured to allow respondents time and opportunity to retrieve pieces of

information that can then be integrated into a more comprehensive judgment of the overall construct. This, they hypothesized, would lead to more reliable judgments, thus reducing undesirable context effects. That is, context effects that are unanticipated or that prompt responses that may be unrelated or only tangentially related to the construct researchers believe they are measuring.

Benton and Daly (1993) conducted interviews to test the hypothesis that a general item about overall satisfaction with city government services would receive less favorable or neutral responses when asked first than if it followed a set of specific questions that asked respondents to consider their satisfaction with specific city services. Results of this study supported their hypothesis. “When the service specific questions were asked first 71.9% of the respondents evaluated the overall quality of services as ‘Very Good’ or ‘Good’ while only 68% did so if the general question preceded the service-specific questions” (Benton & Daly, 1993, p. 641). They concluded that when respondents are put into the position of rendering an overall evaluation without the benefit of time and prompting to first consider the range of municipal services, they are unlikely to make a thoughtful overall appraisal since, as noted earlier, respondents will likely truncate their memory search as soon as they have accessed any relevant attitude or belief about the construct.

Moreover, general items that are also abstract or ambiguous as well as multidimensional, such as an overall rating of satisfaction with the city government,

require respondents to retrieve a larger number of scripts than more concrete general items, such as the overall rating of city services described above. Consequently, general items that are abstract or ambiguous will be even more susceptible to context effects than general items that are concrete. For this reason, multi-item scales of specific items may be used intentionally, as a viable approach, to create context for multidimensional or abstract issues, thus, leading to more stable responses. (Tourangeau, 1987). This approach creates a “salience effect”, a context effect in which “a particular response is made more available or more attractive through a kind of consciousness-raising process created by preceding questions” (Schuman & Presser, 1981, p. 45).

#### Relationships Among Items

Question order effects are most likely to occur when respondents are presented with two or more items about a similar issue. Respondents’ perceptions of the relationships among these items can play a role in determining whether context effects will occur and how they will manifest. In actuality, item relationships may take the form of “part-whole” combinations or “part-part” combinations (Schuman & Presser, 1981). “Part-whole combinations involve two or more questions where one question is more general and is intended to contain, summarize, or imply another in some important sense, but not necessarily in reverse.... Part-part relations involve questions that are at the same level of specificity.... Neither question includes or summarizes the other, though, of course they may have implications for each other” (Schuman & Presser, 198, p. 27).

Frequently, part-part item relationships do not manifest context effects since, as noted earlier, specific items are more resistant to context effects. However, if these part-part context effects do occur, it usually results when the items call for a judgment of fairness, which then evokes a norm of equal treatment (to be discussed later).

Part-whole item relationships.

A part-whole item sequence may produce either consistency or contrast effects depending on the respondent's perceived relationship among the items. A part-whole sequence may produce a consistency effect if the respondent perceives the specific question as being an example to be considered in answering the general question. Conversely, a part-whole contrast effect can occur if the items are perceived as belonging together, but where the general question is asking for information which is different from the specific question (Schuman & Presser, 1981).

A commonly conducted experiment to test for part-whole effects involves asking a question about marital happiness either preceded or followed by a question that asks about general happiness. These studies have produced results ranging from significant consistency effects--to no effects--to significant contrast effects. This points out the vulnerability of this type of item sequencing without understanding and controlling for some of the more common confounding variables. For example, in some studies the specific question immediately precedes or follows the general question, while in others the questions are embedded in a series of other questions. Likewise, in some studies there is a



single specific item related in some way to the general item, while in others there may be a series of related specific items (Strack & Martin, 1987).

Strack and Martin (1987) conducted a study to try to determine under what circumstances part-whole contrast or consistency effects would occur. In one study students were given a two-page questionnaire that contained a question about happiness with dating and a question about general happiness embedded in a series of questions about other student issues. Results showed that when the dating happiness question was asked first, there was a significantly higher correlation between the general and specific happiness questions than when the general happiness question was asked first. In this case, the answer to the general target question was “primed” by the specific question.

In a similar study, a contrast effect was activated by creating a context in which the two questions would be perceived as belonging together. To establish the relationship between the two items, the general and specific happiness questions were asked sequentially and were introduced with the following introductory information: "Now, we would like to learn about two areas of life that may be important for peoples' overall well-being (a) happiness with dating, and (b) happiness with life in general.... As predicted, the correlation between the two measures was significantly reduced" (Strack & Martin, 1987, p 132). In this condition, where the contextual relationship between items was established through introductory material, the general item was interpreted as asking for different information from the specific item.

As just noted, when part-whole contrast effects occur it is because respondents infer that the general question is asking for information that is different from the specific question. In other words, having answered the specific question first, the respondent assumes that the general question is asking for new information, excluding information used to make the earlier specific judgment. This is based on the, so called, “given new contract”, a conversational norm in which the respondent believes he is expected to give new information that is not redundant with earlier responses (Strack & Martin, 1987). This social norm produces a “subtraction effect”, in which, after answering the specific question, the respondent redefines the general question to exclude the specific example (Schuman & Presser, 1981).

One example of this phenomenon occurred in a study conducted by Schuman and Presser (1981) and has been frequently and reliably replicated. Respondents were asked if a pregnant woman should be able “to obtain an abortion if there is a strong chance of a serious defect in the baby” and a question that asks if abortion “should be generally available for married women” (Strack & Martin, 1987, p. 130). This study showed that the proportion of respondents who answered yes to the general question dropped significantly when the specific question was asked first. The answer to the specific question was not affected by first answering the general question. This led researchers to conclude that when the general abortion item was asked first, respondents considered scenarios such as a defective fetus in their judgment process and were, therefore, more

liberal in their support of abortion. However, when the defective fetus item was asked first, respondents excluded this type of scenario from their judgment process and were, therefore, less likely to support the general abortion option (Schuman & Presser, 1981; Strack & Martin, 1987).

#### Part-part item relationships.

Part-part item relationships frequently produce consistency effects when the issue is one that involves a judgment of equal treatment. For example, if respondents answered the first of two questions in a manner that gives special treatment to one group, they will often answer a second related question in the same direction, presumably from a need to show fairness. The most frequently cited example of this effect is an experiment originally conducted in 1948 by Hyman and Sheatsley and replicated by Schuman and Presser in 1980. In this experiment, respondents were asked to answer two questions: (a) One question asked whether respondents thought that the United States should allow Communist reporters from other countries into the United States to report the news as they see it, and (b) The other question asked whether respondents thought a Communist country such as Russia should allow American reporters into Russia to report the news as they see it. Both studies found large, highly significant results which showed that respondents are more likely to say that Communist reporters should be allowed in the United States after they had already answered the question advocating that American reporters be allowed into Russia. Likewise, respondents were less likely to advocate

letting American reporters into Russia after first answering that Communist reporters should not be allowed into the United States. Several other similar types of part-part experiments produced confirming results when the item content invokes “a clear norm of equal treatment for two parties, one of whom may be better liked and therefore (sic) given preferential treatment when the norm is not invoked by context” (Schuman & Presser, 1981, p. 32).

#### Psychological Influences--Attitude Intensity and Issue Involvement

It has been widely accepted by many experienced survey investigators that response effects, including context effects, are, to a large degree, affected by attitude crystallization (Krosnick & Schuman, 1988). Operationally, attitude crystallization has been defined with the constructs attitude intensity and issue involvement. Intuitively, these scientists have reasoned that people who hold strong, well formed opinions on an issue or who are particularly interested in or involved in an issue will be more resistant to the priming effects that sometimes occur from cues present in the survey. The theoretical basis for this assumption is that the well formed attitudes of these types of respondents are associated with more accessible scripts, making them easier to retrieve. Likewise, it is assumed that these well formed attitudes are more resistant to change (Bishop, 1990; Krosnick & Schuman, 1988). On the other hand, many people often have mixed attitudes about issues. That is to say that they can see the value of multiple perspectives; and, as such, they could be swayed either way on the issue. In this case, it is assumed that cues

that are presented in the form of earlier questions will likely influence which of the beliefs that the respondent holds on the issue will be retrieved. For example, if a block of questions about women's rights is placed before a target question about abortion, respondents with mixed beliefs would be more likely to agree that abortion is a matter of a woman's free choice (Tourangeau & Rasinski, 1988). Surprisingly, these assumptions have rarely been tested. Undoubtedly, they are so widely accepted simply because they have such strong intuitive appeal. The few studies that had been conducted yielded mixed results (Krosnick & Schuman, 1988).

Krosnick and Schuman (1988) conducted a study to test the hypothesis that context effects would be smaller among respondents whose attitudes are held with great certainty. In their study, they used the same general-specific abortion item sequence reported earlier. As in the original study conducted by Schuman and Presser (1981), they varied the order of the following two items:

1. Would you tell me whether or not you think it should be possible for a married woman to obtain a legal abortion if she is married and does not want any more children?
2. Please tell whether or not you think it should be possible for a pregnant woman to obtain a legal abortion if there is a strong chance of a serious defect in the baby?

The last question on the survey contained the following text:

The last question is on a different subject. Some people are very certain about their feelings about when legal abortions should be permitted. Other people see this issue as a difficult one to reach a decision on. Would you say that you are more like those who are very certain, or that you are more like those who see this issue as a difficult one to reach a decision on (Krosnick & Schuman, 1988, p. 943)?

The first of these abortion experiments yielded the expected results. Respondents who were very certain about their opinions on legalized abortion were significantly less affected by the order of the two abortion scenarios than respondents who saw abortion as a difficult issue. Three replication attempts, however, failed to reproduce this finding, which led these researchers to conclude that the initial confirming results were likely a chance occurrence.

In a similar type of study, a public opinion survey regarding a number of political issues was conducted to test for the effects of issue involvement on context effects (Bishop, 1990). This study also found no significant relationship between attitude crystallization and context effects. It is important to note, though, that this study did not produce large context effects overall, so there is less possibility to detect the effect of issue involvement.

The most plausible explanation offered for the seemingly counterintuitive results reported in these two studies is that the effects of the attitude intensity or issue

involvement variables may, in fact, have been masked by a context effect. “Because of the usual sequence of the questions (ask about opinion, then involvement) respondents may infer that they are interested in an issue, or that they have thought about it, or that it is important to them because they have just expressed an opinion on it” (Bishop, 1990, p. 217).

### Respondent Characteristics

Some research studies have also looked at the possibility that respondent demographic variables may be interacting with question order to produce context effects. Most have focused on education level, with the hypothesis that less educated respondents will be more prone to context effects because they are not as well informed as more highly educated respondents or, alternatively, because they are not as open-minded (Benton & Daly, 1993; Schuman & Presser, 1981).

Schuman and Presser (1981) looked at the interaction of respondent education and question order in their American-Communist reporter order experiment described earlier, in which the norm of equal treatment led respondents to be more lenient about allowing Communist reporters in the United States after first answering that American reporters should be allowed in Russia. They found that “the effect of order on responses is located mainly among those with less than college experience and tends to disappear among the college educated” (p.31). They attribute this finding to the fact that college educated respondents are much more likely to advocate allowing Communist reporters into the

United States regardless of the order of item presentation, hence, producing a ceiling effect. This finding was reliably replicated from earlier studies applying this same experimental scenario.

Taking this a step further, Benton and Daly (1993) hypothesized that other respondent characteristics, specifically characteristics that may lead respondents to be less informed or less attentive, may lead certain groups of respondents to be more susceptible to context effects. In their study, described earlier, in which respondents were asked a series of questions about satisfaction with specific city services either preceding or following a general question about satisfaction with city services, they looked at the interaction of item order with four demographic variables:

- (1) less educated--those with no formal education through high school complete;
- (2) less attentive--those who reported no contact with city officials in the past year;
- (3) currently employed--those who were still actively employed as opposed to those who were retired; and
- (4) short-term resistant--those who have lived in the city three years or less (p. 500).

Results of this analysis supported their hypothesis. For three of the groups (the less educated, the less attentive, and the currently employed) presenting the general item first produced significantly lower ratings for this item than when the specific items were presented first. The level of significance for short-term residents did not achieve statistical significance, but the relationship was in the hypothesized direction.



### Mode of Administration

Mode of survey administration can have a marked effect on how survey respondents answer attitude and opinion questions. Accordingly, one can expect that context and order effects would behave differently under different survey administration conditions, yet not much research has been done to test these differences. The limited research that has been conducted has shown that, while order effects are more prevalent under telephone or face-to-face interview conditions, other types of context effects are at least as likely, or even more likely to occur under self-administered survey conditions (Schwarz & Hippler, 1995).

To test this hypothesis Schwarz and Hippler (1995) conducted an experiment in which half of the respondents received a mail survey and half were administered the survey by telephone. The target question asked respondents how much money they would be willing to donate to help people suffering in Russia. This question was either immediately preceded or followed by two context questions about taxes. As predicted, the tax questions did have an influence on the amount of money respondents were willing to donate; however, the pattern differed for the two data collection methodologies. In the telephone survey, when the target donation question was asked before the two tax context questions, respondents reported intentions to donate significantly higher sums of money than when the two tax questions were asked first, thus, providing strong support for both context and order effects. However, in the mail survey, order of items did not produce

significant differences in the sum of money respondents intended to donate, but the reported donation intentions were in line with the donation intentions reported in the telephone survey when the tax questions preceded the donation question. This led researchers to conclude that the mail respondents were affected by the context set by the tax questions independent of the order of presentation (Schwarz & Hippler, 1995).

These results can be explained primarily by three factors that differ between telephone or face-to-face interviews and self-administered surveys. First is the sequential or serial order nature of interview questions. In interviews, respondents are held to a strict serial order sequence of the questions. That is, interviewers ask questions in a specified order and respondents are generally either not allowed or are reluctant to go back and change a previous answer based on information that is presented in the form of subsequent survey questions. Thus, if respondents wish to appear consistent, they have no choice but to answer later questions in a manner which is consistent with earlier answers. Conversely, self-administered survey respondents are not bound by this strict serial order presentation; rather, these respondents are afforded the luxury to read ahead or go back and change prior answers at will, based on context set by subsequent questions, which may have rendered a relevant attitude more accessible.

A second noteworthy difference between interview and mail surveys is the element of time pressure. Interview respondents may be more prone to carry-over effects because they feel pressure to come up with answers quickly; and, consequently, they have less time

for the retrieval of relevant information and judgment formation. Accordingly, interview respondents may be more likely to give less thought out answers or, so called, “top of the head” responses (Ayidiya & McClendon, 1990; Schwarz, Strack, Hippler, & Bishop, 1991).

A third notable difference leading to differential responding in different data collection methodologies results from self-selection bias. For example, it is assumed that potential survey respondents with lower education levels are less likely to respond to mail-back surveys because of lower reading proficiency. Additionally, mail respondents are believed to be more interested in the topic because they have the opportunity to preview the survey before deciding to respond. “Given that different modes of data collection do result in markedly different response rates (see Groves and Lyberg, 1988), differential self-selection of respondents along any number of variables provides a plausible explanation for many of the observed mode effects” (Schwarz et al., p. 197).

#### Summary of the Literature

Existing research has established that context effects are more likely to occur in surveys when certain conditions are present. Most notably, when items are general, ambiguous, or abstract respondents are likely to search for contextual cues to provide an interpretive framework. Items that are specific in content are generally resistant to context effects, and multi-item scales consisting of groups of specific items are generally quite stable.

While surveys that ask questions about frequently encountered situations that do not require complex judgments are generally not prone to context effects, attitude accessibility plays an important role in determining which memories a respondent is likely to recall when answering a question. Recent relevant experiences are especially likely to be accessed. Items that are multidimensional, such as asking for an evaluation of some aspect of experience over an extended period of time, are especially vulnerable so asking about a specific experience before a question that calls for an overall evaluation could produce a priming effect. Moreover, research has shown that mail respondents may be affected by context items independent of the order of presentation because they are able to read ahead in search of relevant context.

Furthermore, how respondents interpret their task and the relationships among items can determine whether context effects will occur and whether they will produce either a consistency or contrast effect. When specific and general items are perceived as belonging together, but stems or instructions make clear to respondents that the items are asking for different information, respondents frequently eliminate the specific example from their overall evaluation when the specific items are asked first, thus, producing a contrast effect. Conversely, sometimes respondents will consider the specific item as an example to be considered in answering the general item, thus, producing a consistency effect.

### Background Information on the Current Study

In the increasingly competitive environment of the health care industry, member and patient satisfaction surveys are becoming an integral part of the strategy used to assess operational improvement efforts and to gain market share. In fact, The National Committee on Quality Assurance (NCQA) now mandates the collection of patient satisfaction data as part of their requirement for accreditation. Furthermore, employer groups are increasingly requiring accreditation and member satisfaction guarantees as a condition to offer health plan enrollment to its employees.

To that end, one health maintenance organization designed and implemented a survey tool to assess member and patient satisfaction with aspects of service, access and care. This survey contains a set of 47 general membership items which are introduced with the stem, "Thinking of your experiences over the past twelve months how would you rate ...". Two of these items, which are very general, assess global satisfaction with service and access. The remainder of the general membership items are grouped into multi-item modules relating to overall care experiences, access to care, phones, appointments, and health education experiences over the past twelve months. This survey is sent to low utilizing members who have not had an outpatient visit in the preceding twelve month period as well as to patients who have had a recent outpatient visit. In addition to the general membership items, the outpatient version of the survey has a module of items tacked onto the end that instructs respondents to rate the care they

received from a specific doctor in a specific department on a specific date. Four other small independent surveys are also administered to collect satisfaction data on inpatient experiences and visits to ancillary service departments (pharmacy, laboratory, and radiology).

### Purpose of This Study

In order to reduce the survey burden on its membership and to standardize data collection and reporting across all surveys, this organization wishes to fold its ancillary services modules into the larger general membership survey. Furthermore, there is interest in dropping other custom tailored modules into the general membership survey to assess other specific aspects of performance. Given the propensity for context effects to occur under varying circumstances, it is essential that the effect of potential changes to a survey instrument used as a tracking tool be carefully investigated before implementing these proposed changes.

The purpose of this study is to test whether the changing context brought about by inserting different modules of specific items or reversing the order of the general and specific items will induce respondents to answer these types of questions differently. It is hypothesized that the general, ambiguous items will exhibit context effects, but the modules of specific items will not be affected by context or order of presentation. In line with the prior research on context and order effects, it is expected that

**Hypothesis 1:** Respondents will rate the two global service and access measures and the overall general experience measure differently as a result of order of presentation and service type. Moreover, it is expected that the manifested context effects will be similar for the three ancillary service types (pharmacy, laboratory, and radiology) but different for the outpatient service type.

The context effects are expected to be similar for the three ancillary service types because the ancillary service surveys had very similar service specific items and the same set of general membership items. The context effects are expected to be different for the outpatient service type because the outpatient survey had a greater number of more wide ranging service specific items but fewer general membership items. The literature maintains that mail respondents are just as likely to read ahead to acquire context, if necessary, as they are to use prior items. Consequently, the ancillary service respondents would be provided the same context, while the outpatient respondents would receive different context.

**Hypothesis 2:** Respondents will not rate the visit specific modules differently as a result of order of presentation since specific items are generally resistant to context effect.

Research has shown that specific questions about familiar situations tap readily accessible attitudes. Moreover, items that are specific in nature usually do not require context to help discern the meaning.

**Hypothesis 3:** There will be a stronger relationship between the visit specific overall satisfaction measure and the global service, global access and overall general experience measures when the service specific modules are presented first

Research has shown that when items are multidimensional, such as asking for an evaluation of some aspect of experience over an extended period of time, responses are likely to be primed by specific questions about their most recent experiences. The overall visit satisfaction measure is used as a summary measure of satisfaction with the specific components of the visit because it appears at the end of the service specific modules.

## Method

### Participants

Subjects were health plan members who were chosen at random from records indicating that they had had a recent visit to a doctor or to the pharmacy, laboratory, or radiology department. The total number of surveys administered to the four service groups was 9,300. Total returns was 4,792, yielding an overall response rate of 52%. The number of surveys administered, returned, and the response rates by service type are reported in Table 1. Half of the subjects from the four service types



were randomly selected to receive a version of the survey that presents the general membership items first, and the other half to receive a version that presents the service specific items first. Table 2 presents the sample sizes for each order by service type combination.

### Materials

The survey instruments consisted of eight questionnaires, two for each of the four service types. Each of the questionnaires contained several modules of items assessing general member satisfaction with aspects of care and service over the prior twelve months and a service specific module assessing satisfaction with specific aspects of the visit. The two order versions within service group are identical with the exception that in one version the general items appear first, and in the other version the visit specific items appear first. Most items on all questionnaires use a five-point scale with the anchors Poor, Fair, Good, Very Good, and Excellent. A few of the items in the ancillary service modules use a four-point satisfaction scale with the anchors Very Dissatisfied, Dissatisfied, Satisfied, and Very Satisfied. Additionally, a couple of items on the ancillary service modules have a scale specific to the question such as a cleanliness or comfort scale. A cover letter stating the purpose of the survey and assurance of confidentiality accompanied the questionnaire. Each cover letter was personalized with reference to a specific visit on a specific date to a specific department.

Table 1

Survey Outgo, Returns and Response Rates by Service Type

Service Type	Outgo	Returns	Response Rate
Pharmacy	1,700	804	47%
Laboratory	1,700	898	53%
Radiology	1,700	913	54%
Outpatient	4,200	2,177	52%
Total	9,300	4,792	52%

Table 2

Number of Returns by Order Within Service Type

Service Type	General Items First	Specific Items First	Total
Pharmacy	399 (50%)	405 (50%)	804 (100%)
Laboratory	467 (52%)	431 (48%)	898 (100%)
Radiology	462 (51%)	451 (49%)	913 (100%)
Outpatient	1,100 (51%)	1,077 (49%)	2,177 (100%)
Total	2,428 (51%)	2,364 (49%)	4,792 (100%)

Dependent variables.

Three of the dependent variables used in this study are the second, third, and fourth items of the general measures. All three of these measures are general and multidimensional, but two are also ambiguous. The two items that are general, multidimensional, and ambiguous ask respondents to rate their experiences with “personal and responsive service” and “convenient and easy access”. The item that is general and multidimensional but not ambiguous asks respondents to rate their overall health care experience over the past twelve months.

Three additional dependent variables come from the service specific modules. An overall visit satisfaction measure is used as a summary indicator of satisfaction with the various components of the visit. This measure on the ancillary service surveys, which is stated, “Overall how do you feel about your visit to the [specific department]”, uses the four-point satisfaction scale described earlier. The overall visit satisfaction measure on the outpatient survey, which is stated, “Overall how would you rate this visit to the [specific department]”, uses the five-point poor to excellent scale described earlier. The other two service specific dependent variables are modules of items from each service type that ask respondents to rate wait times and aspects of the care experience. The wait time module consists of two items for the ancillary services and three items for the outpatient service type. The care experience

module consists of nine items for the laboratory and radiology surveys, eight items for the pharmacy survey, and ten items for the outpatient survey.

Context variables.

All material on the survey and cover letter including the visit specific information referenced in the cover letter, survey instructions and stems, as well as all survey items serve as context variables that change as a result of varying the service type and order of presentation. There were two important factors contributing to the differing context between the ancillary services and the outpatient service type:

- (a) The general membership modules for the ancillary service surveys are identical, but the outpatient survey contains fewer items in each of the general modules, and
- (b) The specific modules for the three ancillary service surveys vary only slightly in length and content, but the outpatient module is longer and more detailed, with several different types of additional service specific items.

In total, the general membership modules consist of 47 items on the ancillary service surveys and 28 items on the outpatient survey. All general membership items use the five-point poor to excellent scale. The general modules for all surveys begin with a global assessment of quality health care, followed by the two global service and access items that serve as dependent variables. Below that, in order of presentation, are modules that assess satisfaction with aspects of the general care experience, access to care, phone service, appointment availability, and self-care advice and information.

The “general” module for all service types includes the overall health care experience item, which serves as a dependent variable, as well as items addressing respect for members’ time, coordination among caregivers, follow-up after a medical problem, and manner in which test results are conveyed. The ancillary surveys also contain additional items in this module addressing courtesy, helpfulness, and caring attitude of physicians and staff and the extent to which respondents were involved in decisions about their medical care.

The “access to care” module for all service types includes items assessing assistance provided and ease in selecting a personal provider, ability to see a health care provider when needed, ability to see a particular provider, and ability to get after hours care (the after hours care assessment consists of three items for the ancillary service surveys but is assessed by one item on the outpatient survey). In the ancillary surveys, the block of access items just described is preceded by items assessing responsiveness to phone calls, wait time at the medical facility, and amount of time providers spend with them and is followed by items assessing ability to get timely specialty consults.

The “phones” module for all service types assesses respondents’ perceptions of their ability to get through on the phone to an appropriate problem-solver at a convenient time, helpfulness of phone advice received, and availability of after hours phone advice (the after hours phone assessment consists of three items for the ancillary surveys but is assessed by one item on the outpatient survey). For the outpatient service type, the

phones module also contains the responsiveness to phone calls item which appears in the access to care module on the ancillary service surveys. The ancillary surveys contain an additional item assessing courtesy and helpfulness of phone staff.

The “appointments” module differs the most between the ancillary surveys and the outpatient survey. For the outpatient survey, this module asks just three questions assessing wait times for an urgent care, a non-urgent, or a return appointment. The ancillary surveys include items about ability to get an appointment and convenience of appointment time, as well as four items assessing wait times for appointments involving four rather detailed scenarios.

For the “keeping you healthy module”, which is the last of the general modules, all surveys present items assessing how well their health care provider teaches them to care for common problems, their confidence in knowing when self-care is appropriate, how well they are informed when preventive tests are needed, and ability to get advice on specific actions to stay healthy. The ancillary surveys contain two additional items assessing ability to get services and advice to help stay healthy and ability to get preventive tests.

The service specific modules all have in common items that assess satisfaction with wait times and the care experience. These are the two modules that also serve as dependent variables. As noted earlier, the ancillary surveys all contain two items assessing satisfaction with wait times (these items use the four-point satisfaction scale), while the

outpatient survey has three wait time items (these items use the five-point poor to excellent scale). The care experience modules for the ancillary service surveys, while not identical, are very similar in content. These items assess aspects of the care experience such as courtesy, treatment, and professional appearance of professional and staff personnel; information, explanations, and instructions; and professional skills or performance. The care experience module on the outpatient survey contains items that are more detailed and specific. Many of these items assess aspects of the provider-patient interaction: Was the provider easy to talk to; a good listener; courteous, caring and respectful? Did the provider spend enough time with you, involve you in medical decisions, explain things in terms you could understand, adequately answer your questions? There are also two questions assessing the provider's professional credibility: (a) familiarity with medical history and (b) the provider's skills and abilities. All care experience items on all surveys utilize the five-point poor to excellent scale. Beyond the two modules just described, the ancillary service surveys contain items assessing the comfort and cleanliness of the department and one or two overall satisfaction measures. The outpatient survey, however, has several additional items assessing aspects of scheduling the appointment, appointment convenience, familiarity with the provider, and access to specialty care, as well as an overall visit satisfaction measure.



### Procedure

Surveys were mailed by an independent vendor to subjects approximately one to two months after their visit with a cover letter explaining the purpose of the survey and a self-addressed stamped envelope. It should be noted that there was a delay in mailing the ancillary service surveys, hence, the outpatient surveys were mailed approximately one month after the visit occurred, while the ancillary service surveys were mailed approximately two months after the visit. The data collection cycle lasted for ten weeks with extensive follow-up procedures to ensure the highest response rates possible. The follow-up consisted of a reminder postcard mailed to non-respondents two weeks after the initial survey, a second copy of the survey mailed two weeks after the postcard and finally a phone call two weeks after the second survey.

### Results

#### Effects of Order and Service Type on General Items

The first hypothesis consists of two parts: (a) There will be an effect of order and service type for the three general measures--global service, global access, and overall general satisfaction, and (b) The effect will be similar for the three ancillary service types but different for the outpatient service type. Since a consistent effect was expected for the three ancillary services, these service types were tested first; so that, if they were not significantly different from each other they could be combined for the comparison with the outpatient service type. To test the hypothesis that respondents will rate the two global

service and access measures and the overall general experience measure similarly for the three ancillary service surveys, three 2 x 3 analyses of variance (ANOVAs) were computed to test for a nonsignificant main effect of service type and for a nonsignificant interaction of order and service type for the pharmacy, laboratory, and radiology service types (see Table 3 for descriptive statistics).

The first ANOVA tested for the main effect of service type and for the interaction between order and service type for the three ancillary services on the global service measure. As predicted there was not a significant main effect for service type  $F(2, 2457) = .19$ . Likewise, there was not a significant interaction between order and service type on this measure  $F(2, 2457) = .25$  (see Table 4). The second ANOVA tested for the main effect of service type and the interaction between order and service type for the three ancillary services on the global access measure. Consistent with the results above, there was not a significant main effect for service type for this measure either  $F(2, 2429) = .11$ , nor was there a significant interaction between order and service type  $F(2, 2429) = .67$  (see Table 5). Finally, the third ANOVA tested for the main effect of service type and for the interaction of order and service type for the three ancillary services on the overall general experience measure. Once again, as predicted, there was not a significant main effect for service type  $F(2, 2492) = .14$ , nor was there a significant interaction between order and service type  $F(2, 2492) = 1.21$  (see Table 6).

Given that no significant main effects or interactions were found among the three ancillary service types, these service types were combined to test for the effects of order and service type between the ancillary services and the outpatient service type. Three 2 x 2 ANOVAs were conducted to test the hypothesis that there will be a significant effect of order and service type between the ancillary services and the outpatient service type for the global service and access measures and for the overall general experience measure (see Table 3 for descriptive statistics).

The first ANOVA tested for an interaction between order and service type for the global service measure. As predicted, a significant interaction was found  $F(1, 4514) = 5.20, p < .05$  (see Table 4). One-way ANOVAs were conducted, as a follow-up to the significant interaction, to test the simple main effects of service type when controlling for order of presentation. When the general membership items were presented first, the outpatient respondents rated this item significantly higher ( $M = 3.73$ ) than the ancillary service respondents ( $M = 3.63$ ),  $F(1, 4514) = 5.20, p < .05$ . However, when the service specific items were presented first, there was not a significant difference between the service types  $F(1, 4514) = .91$ . One-way ANOVAs were also conducted to test for the simple main effects of order when controlling for service type. For the ancillary services a significant order effect did not occur  $F(1, 4514) = 2.06$ . Likewise there was not a significant order effect for the outpatient service type on this measure  $F(1, 4514) = 3.16$ . However, in contrast to the effect that resulted when the general items were presented

first, a crossover effect occurred when the service specific items were presented first, with the mean ratings for the outpatient group ( $\underline{M} = 3.65$ ) dropping below the ancillary service mean ( $\underline{M} = 3.69$ ), thus, producing the interaction.

The second ANOVA tested for the interaction between order and service type for the global access measure. Once again the hypothesis was supported with a significant interaction  $\underline{F}(1, 4440) = 10.04, p < .01$  (see Table 5). One-way ANOVAs were conducted, as a follow-up to the significant interaction, to test the simple main effects of service type when controlling for order of presentation. In this case, when the general membership items were presented first, there was not a significant difference between the ancillary services and the outpatient service type  $\underline{F}(1, 4440) = 1.74$ . However, when the service specific items were presented first, the ancillary service respondents rated this measure significantly higher ( $\underline{M} = 3.73$ ) than the outpatient respondents ( $\underline{M} = 3.59$ ),  $\underline{F}(1, 4440) = 9.96, p < .01$ . One-way ANOVAs were conducted to test for the simple main effects of order when controlling for service type. In this case, the ancillary service respondents rated this measure significantly lower when the general items were presented first ( $\underline{M} = 3.64$ ) than when the service specific items were presented first ( $\underline{M} = 3.73$ ),  $\underline{F}(1, 4440) = 4.03, p < .05$ . Conversely, the outpatient respondents rated this measure significantly higher when the general items were presented first ( $\underline{M} = 3.70$ ) than when then service specific items were presented first ( $\underline{M} = 3.59$ ),  $\underline{F}(1, 4440) = 6.07, p < .05$ .

Table 3

Descriptive Statistics For General Dependent Measures

<b>Global Service Measure</b>	<b>General Items First</b>			<b>Specific Items First</b>		
<b>Service Type</b>	<b><u>M</u></b>	<b><u>SD</u></b>	<b><u>n</u></b>	<b><u>M</u></b>	<b><u>SD</u></b>	<b><u>n</u></b>
Pharmacy	3.63	.96	375	3.69	1.07	388
Laboratory	3.60	1.03	447	3.70	1.06	410
Radiology	3.67	1.04	413	3.69	1.07	430
Combined Ancillaries	3.63	1.01	1235	3.69	1.06	1228
Outpatient	3.73	1.00	1046	3.65	1.06	1009
<b>Global Access Measure</b>	<b>General Items First</b>			<b>Specific Items First</b>		
<b>Service Type</b>	<b><u>M</u></b>	<b><u>SD</u></b>	<b><u>n</u></b>	<b><u>M</u></b>	<b><u>SD</u></b>	<b><u>n</u></b>
Pharmacy	3.65	1.04	373	3.74	1.06	384
Laboratory	3.60	1.07	445	3.75	1.01	401
Radiology	3.68	1.10	411	3.70	1.11	421
Combined Ancillaries	3.64	1.07	1229	3.73	1.06	1206
Outpatient	3.70	1.06	1027	3.59	1.09	982

Table 3 (continued)

Descriptive Statistics For General Dependent Measures

Overall Experience Measure	General Items First			Specific Items First		
	<u>M</u>	<u>SD</u>	<u>n</u>	<u>M</u>	<u>SD</u>	<u>n</u>
Pharmacy	3.78	.92	386	3.78	.99	392
Laboratory	3.70	.93	448	3.81	.98	417
Radiology	3.78	.97	419	3.75	1.02	436
Combined Ancillaries	3.75	.94	1253	3.78	1.00	1245
Outpatient	3.80	.94	1068	3.75	.97	1026

Table 4

Analysis of Variance for Global Service Measure

<b>Factorial ANOVAs</b>	<b>Three Ancillary Services<sup>a</sup></b>		<b>Ancillaries vs Outpatient</b>	
Source	<u>df</u>	<u>F</u>	<u>df</u>	<u>F</u>
Service Type	2	.19	1	.86
Order	1	1.97	1	.12
Service Type x Order	2	.25	1	5.20*
Error	4514	(1.08)	4514	(1.07)
<b>One-Way ANOVAs</b>	<b>General Items First</b>		<b>Specific Items First</b>	
Source	<u>df</u>	<u>F</u>	<u>df</u>	<u>F</u>
Service type within order	1	5.20*	1	.91
Error	4514	(1.07)	4514	(1.07)
<b>One-Way ANOVAs</b>	<b>Combined Ancillaries</b>		<b>Outpatient Service</b>	
Source	<u>df</u>	<u>F</u>	<u>df</u>	<u>F</u>
Order within service type	1	2.06	1	3.16
Error	4514	(1.07)	4514	(1.07)

Note. Values enclosed in parentheses represent mean square errors.

<sup>a</sup>In the absence of an interaction or a significant main effect for service type in the first factorial ANOVA, the three ancillary services were combined for subsequent analyses.

\* $p < .05$ . \*\* $p < .01$ .

Table 5

Analysis of Variance for Global Access Measure

<b>Factorial ANOVAs</b>	<b>Three Ancillary Services<sup>a</sup></b>		<b>Ancillaries vs Outpatient</b>	
Source	<u>df</u>	<u>F</u>	<u>df</u>	<u>F</u>
Service Type	2	.11	1	1.75
Order	1	3.98*	1	.22
Service Type x Order	2	.67	1	10.04**
Error	2429	(1.13)	4440	(1.14)
<b>One-Way ANOVAs</b>	<b>General Items First</b>		<b>Specific Items First</b>	
Source	<u>df</u>	<u>F</u>	<u>df</u>	<u>F</u>
Service type within order	1	1.74	1	9.96**
Error	4440	(1.14)	4440	(1.14)
<b>One-Way ANOVAs</b>	<b>Combined Ancillaries</b>		<b>Outpatient Service</b>	
Source	<u>df</u>	<u>F</u>	<u>df</u>	<u>F</u>
Order within service type	1	4.03*	1	6.07*
Error	4440	(1.14)	4440	(1.14)

Note. Values enclosed in parentheses represent mean square errors.

<sup>a</sup>In the absence of an interaction or a significant main effect for service type in the first factorial ANOVA, the three ancillary services were combined for subsequent analyses.

\* $p < .05$ . \*\* $p < .01$ .



Table 6

Analysis of Variance for Overall Experience Measure

Source	Three Ancillary Services <sup>a</sup>		Ancillaries vs Outpatient	
	<u>df</u>	<u>F</u>	<u>df</u>	<u>F</u>
Service Type	2	.14	1	.27
Order	1	.46	1	.17
Service Type x Order	2	1.21	1	1.97
Error	2492	(.94)	4588	(.92)

Note. Values enclosed in parentheses represent mean square errors.

<sup>a</sup>In the absence of an interaction or a significant main effect for service type in the first factorial ANOVA, the three ancillary services were combined for subsequent analyses.

Finally, the third ANOVA tested for the interaction between order and service type for the overall experience measure. In this analysis a significant interaction was not found  $F(1, 4588) = 1.97$  (see Table 6).

#### Effect of Order on Service Specific Modules

To test the hypothesis that the service specific modules will be resistant to the context effects created by order reversal, one-way ANOVAs were computed to test for significant differences resulting from order of presentation for two different service specific modules (wait times and care experiences) for each of the four service types. In this case, the service types were tested separately for order effects because each service type survey is measuring a different type of service experience; therefore, it would not be expected that the ratings for the service specific experiences would be comparable (see Table 7 for descriptive statistics).

As predicted, results of the ANOVAs that tested for the effects of order on the wait time modules showed that there were no significant differences in ratings of wait times for the pharmacy service type  $F(1, 723) = .38$ , the laboratory service type  $F(1, 854) = .27$ , the radiology service type  $F(1, 852) = .49$ , or the outpatient service type  $F(1, 2131) = 0.00$  (see Table 8). Likewise, results of the ANOVAs that tested for the effects of order on the care experience modules showed that there were no significant differences in ratings of care experiences for the pharmacy service type  $F(1, 776) = .17$ , the radiology service type  $F(1, 867) = .18$ , or the outpatient service type  $F(1, 2115) = .14$ . However,

for the laboratory service type, respondents rated their satisfaction with their care experience significantly higher when the service specific items were presented first  $F(1, 874) = 4.36, p < .05$  (see Table 8).

#### Effect of Order on the Relationship of Overall Visit Satisfaction and General Measures

To test the hypothesis that there will be a stronger relationship between the overall visit satisfaction measure and the global service, global access, and overall general care experience measures when the service specific items are presented first, correlations were computed for each general-specific item combination for each service type and order condition. Z-tests were conducted to test for statistically significant differences in correlations resulting from order of presentation.

As shown in Table 9, the correlation between the overall visit satisfaction measure and the global service measure was significantly higher for the pharmacy group when the service specific items were presented first  $z(1, 620) = 3.06, p < .01$  and for the laboratory group  $z(1, 703) = 2.01, p < .05$ . The correlations were not significantly different for the radiology group  $z(1, 593) = 1.11$  or for the outpatient group  $z(1, 1992) = 1.80$ .

The correlation between the overall visit satisfaction measure and the global access measure was significantly higher for the outpatient group when the service specific items were presented first  $z(1, 1954) = 3.18, p < .01$ . The correlations were not significantly different for the pharmacy group  $z(1, 613) = 1.90$ , the laboratory group  $z(1, 693) = 1.65$ , or the radiology group  $z(1, 591) = .36$ .

Table 7

Descriptive Statistics For Specific Dependent Measures

<b>Wait Times Module</b>	<b>General Items First</b>			<b>Specific Items First</b>		
<b>Service Type</b>	<b><u>M</u></b>	<b><u>SD</u></b>	<b><u>n</u></b>	<b><u>M</u></b>	<b><u>SD</u></b>	<b><u>n</u></b>
Pharmacy	3.33	.64	359	3.36	.63	366
Laboratory	3.20	.65	440	3.22	.61	416
Radiology	3.34	.60	421	3.31	.62	433
Outpatient	3.58	.92	1081	3.58	.96	1052
<b>Care Experience Module</b>	<b>General Items First</b>			<b>Specific Items First</b>		
<b>Service Type</b>	<b><u>M</u></b>	<b><u>SD</u></b>	<b><u>n</u></b>	<b><u>M</u></b>	<b><u>SD</u></b>	<b><u>n</u></b>
Pharmacy	3.75	.86	384	3.78	.87	394
Laboratory	3.75	.87	453	3.87	.83	423
Radiology	3.89	.86	427	3.92	.84	442
Outpatient	4.13	.92	1063	4.12	.92	1054

Note. Because care experiences differ among service types, it was not expected that ratings across service types would be consistent. Therefore, only the order condition was tested for these measures.

Table 8

One-Way Analysis of Variance for Service Specific Modules

<b>Pharmacy</b>	Wait Times Module		Care Experience Module	
Source	<u>df</u>	<u>F</u>	<u>df</u>	<u>F</u>
Order Within Service Type	1	.38	1	.17
Error	723	(.41)	776	(.75)
<b>Laboratory</b>	Wait Times Module		Care Experience Module	
Source	<u>df</u>	<u>F</u>	<u>df</u>	<u>F</u>
Order Within Service Type	1	.27	1	4.36*
Error	854	(.40)	874	(.73)
<b>Radiology</b>	Wait Times Module		Care Experience Module	
Source	<u>df</u>	<u>F</u>	<u>df</u>	<u>F</u>
Order Within Service Type	1	.49	1	.18
Error	852	(.37)	867	(.72)
<b>Outpatient</b>	Wait Times Module		Care Experience Module	
Source	<u>df</u>	<u>F</u>	<u>df</u>	<u>F</u>
Order Within Service Type	1	.00	1	.14
Error	2131	(.89)	2115	(.84)

Note. Values enclosed in parentheses represent mean square errors.

\* $p < .05$ .

Table 9

Correlations of Overall Visit Satisfaction With General Dependent Measures

Global Service Measure	General Items First		Visit Items First		Order Difference
	r	n	r	n	z
Pharmacy	.26	304	.47	316	3.06**
Laboratory	.29	359	.42	344	2.01*
Radiology	.35	287	.43	306	1.11
Outpatient	.45	1010	.51	982	1.80
Global Access Measure	r	n	r	n	z
Pharmacy	.31	301	.44	312	1.90
Laboratory	.28	357	.40	336	1.65
Radiology	.37	288	.40	303	.36
Outpatient	.30	995	.42	959	3.18**
Overall Experience Measure	r	n	r	n	z
Pharmacy	.36	312	.47	319	1.71
Laboratory	.29	363	.44	349	2.27*
Radiology	.39	288	.51	310	1.85
Outpatient	.44	1032	.52	1002	2.36*

Note. All correlations are significant at the .001 level.

\*p < .05. \*\* p < .01.

The correlation between the overall visit satisfaction measure and the overall general care experience measure was significantly higher for the outpatient group when the service specific items were presented first  $z(1, 2034) = 2.36, p < .05$ . Likewise, for the laboratory group, the correlation between these two measures was significantly higher when the service specific items were presented first  $z(1, 712) = 2.27, p < .05$ . The difference in correlations for the pharmacy group was not statistically significant when the service specific items were presented first  $z(1, 631) = 1.71$ , nor was the difference significant for the radiology group  $z(1, 598) = 1.85$ . In all cases, the correlations were higher when the service specific items were presented first even though only about half of the comparisons reached significance. All correlations for both order conditions were statistically significant ( $p < .001$ ) (see Table 9).

#### Discussion

The purpose of this study was to test whether the changing context brought about by inserting different modules of specific items or reversing the order of general and specific items would induce respondents to answer these general and specific items differently. This section will discuss, in more detail, the findings and implications of this study, as well as the strengths and limitations and suggestions for future research.

The first hypothesis predicted that the global service and access measures and the overall general experience measure would manifest context effects brought about by including modules of specific items about different types of service experiences as well as

by changing the order of presentation because these items are, to varying degrees, general, ambiguous and multidimensional. Furthermore, it was predicted that the effect of context on these three measures would manifest similarly for the three ancillary service types (pharmacy, laboratory, and radiology) because these three service types had very similar service specific modules and the same general membership items, thus, providing respondents with very similar context. On the other hand, it was predicted that the effect of context for these three general measures would be different for the outpatient service type because this service type had a larger number of more wide ranging service specific items and fewer general membership items, thus, providing a substantially different context. These hypotheses were partially supported. As predicted, the ancillary service surveys did, in fact, behave very similarly for these three measures. There were no significant main effects or interactions among these three service types; hence, they were treated as one service type for subsequent order by service type analyses.

The three dependent variables (global service, global access, and the overall general experience measure) are all general and multidimensional, but they vary in degree of ambiguity. Interestingly, and not entirely unexpected, the degree of ambiguity appears to have played an important role in the degree to which context effects were manifested. As noted earlier, when faced with an ambiguous item, respondents are especially likely to search for contextual cues to serve as an interpretive framework (Tourangeau, 1987; Tourangeau & Rasinski, 1988). The item measuring satisfaction with one's overall



experiences, while certainly multidimensional, appears to be the clearest and most straightforward of the three general dependent measures. This item did not yield any significant differences as a result of service type or order of presentation.

The item that measures personal and responsive service could be judged as more ambiguous since, while it is a concept that is easily understood, service can have a multitude of meanings, from respect for your time to kindness and courtesy, for example. In this case, respondents may try to determine what the item means in the context in which it is presented. This item did yield a statistically significant difference between the ancillary services and the outpatient service type, but only when the general membership items were presented first. There was not a significant difference between these two service types when the service specific items came first, nor were there significant differences for either of the service types resulting from order of presentation.

The most likely explanation for the seemingly curious disparity in ratings between the ancillary and outpatient service types on the global service measure when the general items are presented first is that the ancillary surveys have considerably more general membership items than appear on the outpatient survey. Since these items appear after the global service item, this result implies that respondents on self-administered surveys may, in fact, read ahead to find context for general, ambiguous items (Schwarz & Hippler, 1995). This seems an especially likely explanation because the ancillary services did not differ from each other on this item. Items that follow the global service item closely on

the ancillary service surveys but do not appear on the outpatient survey are related to courtesy, helpfulness, and caring attitude of providers and staff.

Another possible intervening factor may have been the combination of the cover letter and the delayed mailing of the ancillary service surveys for a full month after the mailing of the outpatient surveys. Cover letters for all service types primed respondents that they would be asked to answer questions about their general experiences as well as questions about a specific referenced visit that happened on a specific date and in a specific department. It is not clear what the impact of the two month delay between the visit and mailing of the ancillary service surveys might be, but one can speculate that respondents who received these delayed surveys would be suffering considerably more memory deterioration than the outpatient responders who had only a one month gap between the visit date and receipt of the survey. This could render outpatient responders more likely to reference their recent outpatient experience for context when answering the general service and access measures, while ancillary responders are left to search for contextual cues elsewhere.

Additionally, it seems reasonable to assume that most respondents who utilized one of the ancillary services most likely also had a concurrent outpatient visit in conjunction with the laboratory, radiology, or pharmacy experience since a referral or a prescription are generally required to utilize these services. This adds additional

complexity to the possible interaction of factors leading to the disparity in responses between the ancillary and outpatient respondents.

The item which measures convenient and easy access could also be judged as quite ambiguous, presenting respondents with the task of deciding whether this concept means close proximity to their home, plenty of parking, public transportation and handicapped ramps or ability to get appointments, get through on the phone, or be referred to specialists. This item yielded significant differences between the two service types when the service specific items were presented first but not when the general membership items were presented first. Additionally, both the ancillary and outpatient service types yielded statistically significant differences for this item resulting from order of presentation. However, the ancillary services showed an increase in ratings when the service specific items came first, while the outpatient service type showed a decrease in ratings when the service specific items came first. The global service item exhibited this same crossover pattern, but the disparity between the service groups was significantly different when the general membership items were presented first but not when the service specific items were presented first. Conversely, for the access item, the disparity between the service groups was significantly greater when the service specific items were presented first but not when the general membership items were presented first.

The smaller disparity between the ancillary and the outpatient service types for the global access measure when the general items are presented first, as compared to the

global service results, could be due to the ambiguity of the global access measure. While outpatient responders may have been primed by the cover letter to consider their recent outpatient experience for the global service measure, it is unlikely that this overall memory impression could have provided enough context to disambiguate the global access measure. If this is the case, all respondents would be equally likely to search for contextual cues that appear in later survey items. It is unlikely, however, that most respondents would read ahead to the service specific items since there are many items between the global items and the service specific modules when the general items come first. The literature maintains that respondents are likely to truncate their search as soon as any relevant information has been accessed (Benton & Daly, 1993; Bishop, 1987).

For all service types, the second general membership module appearing after the global access measure is prominently labeled "Access To Care". The first several items in this module are the same for all service types. If, indeed, respondents are reading ahead for contextual cues then all service types would be considering similar dimensions of access. This could then lead to the salience effect described earlier by Schuman and Presser (1981) in which specific items are perceived as examples to be considered in answering a general item and, therefore, are used as context for an overall evaluation item.

When the service specific items are presented first a different pattern emerges. The ancillary service ratings of the global access measure increase significantly while the outpatient ratings of this measure decrease significantly, clearly indicating that the service

specific items are influencing global access ratings differentially. This is an especially likely explanation because when the service specific items are presented first, the global items immediately follow. Apart from being influenced by specific aspects of the visit, obscured from the general memory impression of responders who answered the general items first, one can speculate that ancillary service responders, who use the visit as context to disambiguate the item, may have enhanced access perceptions by virtue of the fact that they were, in fact, referred to one or more of these “extra” services, thus, increasing their satisfaction ratings.

The most plausible explanation for the decrease in outpatient ratings for the global access item when the service specific items are presented first is that there are several items addressing specific aspects of access on the outpatient module that do not appear in the ancillary service specific modules. Two of these items, which measure satisfaction with access to specialty care, appear just before the global measures with only two other items between them. These items most likely had a priming effect for the global access measure. A follow-up correlation analysis to test this theory did, in fact, show that the relationships between these two access to specialty care items and the global access item were significantly higher when the service specific module was presented first.

The second hypothesis predicted that respondents would not rate the service specific modules differently as a result of order of presentation. This hypothesis was, for the most part, supported. The service specific items were separated into two modules for

each service type. One module rated wait times and the other rated the care experience. Each service type was tested separately for order effects. All but one of the order comparisons produced nonsignificant differences, thus, supporting the hypothesis. The laboratory was the one notable exception showing a significantly higher rating for the care experience when the service specific items were presented first. Since this was the only measure out of eight showing a statistically significant difference, it can probably be attributed to chance.

The third hypothesis predicted that the correlation between the overall visit satisfaction measure, which appeared in each of the service specific modules, and the global service, global access, and overall general experience measures would be stronger when the service specific items appeared first. The service specific satisfaction measure was used as a summary indicator of satisfaction with the various components of the visit since it appeared at the end of the service modules; hence, in line with the literature, it would most likely provide a stable summary measure of service specific satisfaction. This hypothesis received limited support. Though all of the differences in correlations were in the predicted direction, not all of these differences achieved statistical significance. The correlation between the overall visit satisfaction and the overall experience measures was significantly higher when the service items came first, for the outpatient and laboratory service types but not for the pharmacy and radiology. Likewise, the correlation between the overall visit satisfaction measure and the global service measure was significantly

higher for the pharmacy and laboratory but not for the outpatient or radiology departments. Finally, the correlation between the overall visit satisfaction measure and the global access measure was significantly higher for the outpatient service type but not for the three ancillary services.

In all cases, the correlations between the overall visit satisfaction measure and the global service and access measures were higher when the service specific items were presented first, even though for some groups the difference did not reach statistical significance. This may be due, in part, to the fact that all correlations were highly significant for both order conditions. It is also worth noting that when comparing the magnitude of the correlations between service types, there is more disparity when the general items are presented first. A far more consistent pattern emerges when the service specific items are presented first, lending further credibility to the hypothesis.

#### Implications of This Study

Evidence that supports the presence of context and order effects emerged in this study. Many of the results were consistent with the literature, but not all hypotheses were fully supported. This is also in line with the literature which asserts that the occurrence of context effects is frequently difficult to predict (Krosnick & Alwin, 1987). This is true because, as has been delineated throughout the literature review, there are a host of factors other than the generalness or specificity of the items that may predispose respondents to utilize contextual cues. This study has focused on the general or specific

nature of the target items as the primary predictor of context effects; but respondent characteristics such as length of membership, frequency of utilization, or education may also be exerting an influence and, as such, may merit further research.

Despite the possible influence of other intervening factors, the fact remains that this study did provide evidence that items that are general, multidimensional, and ambiguous are prone to context effects, while items that are specific are generally resistant to context effects. This has implications for survey designers and for organizations using surveys to make important policy decisions. It is safer to use specific items or, better yet, a multi-item scale of specific items to measure a construct of interest. If an overall indicator of the construct is needed then one can use either the scale average or include the overall measure after the block of specific items. In this way the questionnaire can lead respondents through the process of accessing the relevant attitudes necessary for rendering a thoughtful overall evaluation.

### Strengths and Limitations

The main strengths of this study are related to the fact that it was an actual experiment conducted by this health care organization for the purpose of improving their research methodology. As such, respondents are “real” health plan members who had “real” health care experiences. There was nothing about the treatment of the subjects that was artificial, manipulated or contrived. The surveys were mailed, as always, by the independent vendor who regularly administers the ongoing surveys. This limits many of



the confounding influences present in many research studies. Other strengths were that there were large sample sizes, good response rates, and a variety of treatment conditions. One of the obvious weaknesses in the study was the time lag between the visits and the survey mailing. As noted earlier, this is especially problematic because the time lag was not consistent across all service types--a one month delay for the outpatient surveys and a two month delay for the ancillary service surveys. This makes it difficult to determine whether between group differences were due to this inconsistency or to true effects.

Another weakness was the inconsistency in the number of the general membership items between the ancillary service surveys and the outpatient surveys. As noted earlier, the outpatient surveys contained 28 of the general membership items, while the ancillary service surveys each contained 47 such items. This provides another confounding influence which limits one's ability to explain or generalize about the results.

#### Future Research

Regarding future research, the author suggests conducting a similar order by service type experiment but keeping the number and content of general items consistent across all service types. This would allow a cleaner interpretation of the effects that occur. Additionally, a similar study that could achieve sample sizes large enough to analyze the results by length of membership and frequency of utilization could provide insight into the affect of these variables on a respondent's propensity or resistance to context effects.

It seems of critical importance for utilizers of survey data to have a clear understanding of what respondents think they are rating. To the extent that there is a consistent set of perceptions between respondents and researchers about what is being asked and what aspects of the construct should be considered in answering the question then, to that extent, decision-makers will have valid results. It seems clear from the results of this and other studies that the best approach, from a validity standpoint, to measuring respondent attitudes about a given construct is to include multi-item scales of specific items. In this way, context effects can be used to good advantage by cueing respondents about what to consider when providing an overall evaluation. Unfortunately, the practicalities of survey administration, such as the costs involved in the extensive follow-up frequently necessary to get acceptable response rates, often preclude the use of lengthy surveys. In such cases, survey designers should carefully weigh which discrete constructs are of most critical importance and then limit the survey to these critical few. Then a set of specific items about each of these constructs should be included.

## References

- Ayidiya, S. A. & McClendon, M. J. (1990, Summer). Response effects in mail surveys. Public Opinion Quarterly, *54* (2), 229-247.
- Benton, J. E. & Daly, J. L. (1993, Winter). Measuring citizen evaluations: the question of question order effects. Public Administration Quarterly, *16* (4), 492-507.
- Bishop, G. F. (1990, Summer). Issue involvement and response effects in public opinion surveys. Public Opinion Quarterly, *54* (2), 209-218.
- Harrison, D. A. & McLaughlin, M. E. (1993). Cognitive processes in self-report responses: tests of item context effects in work attitude measures. Journal of Applied Psychology, *78*, 129-140.
- Krosnick, J. A. & Alwin, D. F. (1987). An evaluation of a cognitive theory of response order effects in survey measurement. Public Opinion Quarterly, *51*, 201-219.
- Krosnick, J. A. & Schuman, H. (1988). Attitude intensity, importance, and certainty and susceptibility to response effects. Journal of Personality and Social Psychology, *6*, 940-952.
- Schuman, H. & Presser, S. (1981). Questions and answers in attitude surveys. New York: Academic Press.
- Schwarz, N. & Hippler, H. J. (1995). Subsequent questions may influence answers to preceding questions in mail surveys. Public Opinion Quarterly, *59*, 93-97.
- Schwarz, N., Strack, F., Hippler, H. J., & Bishop, G. (1991). The impact of administration mode on response effects in survey measurement. Applied Cognitive Psychology, *5*, 193-212.
- Strack, F. & Martin, L. (1987). Thinking, judging, and communicating: a process account of context effects in attitude surveys. In H. J. Hippler, N. Schwarz, & S. Sudman (Eds.), Social information processing and survey methodology (pp. 123-148). New York: Springer-Verlag.
- Teas, R. K. & Wong, J. K. (1992, March). Item context and the stability of entity-based and attribute-based multiattribute scaling methods. Journal of Consumer Research, *18*, 536-545.

Tourangeau, R. (1987). Attitude measurement: a cognitive perspective. In H. J. Hippler, N. Schwarz, & S. Sudman (Eds.), Social information processing and survey methodology (pp. 149-162). New York: Springer-Verlag.

Tourangeau, R. & Rasinski, K. A. (1988). Cognitive processes underlying context effects in attitude measurement. Psychological Bulletin, 3, 299- 314.

RECD JUL 07 1997



**San José State**  
UNIVERSITY

**Office of the Academic  
Vice President**  
**Associate Vice President**  
**Graduate Studies and Research**

One Washington Square  
San José, CA 95192-0025  
Voice: 408-924-2480  
Fax: 408-924-2477  
E-mail: [gstudies@wahoo.sjsu.edu](mailto:gstudies@wahoo.sjsu.edu)  
<http://www.sjsu.edu>

TO: Susan Rosenberg  
3090 Balmoral Dr.  
San Jose, CA 95132

FROM: Adrian Rodriguez, Ph.D. *Adrian Rodriguez*  
Chairperson, Human Subjects Review Board

DATE: June 30, 1997

The Human Subjects-Institutional Review Board has approved your request to use human subjects in the study entitled:

"Context and Order Effects in  
Self-Administered Surveys"

This approval is contingent upon the subjects participating in your research project being appropriately protected from risk. This includes the protection of the anonymity of the subjects' identity when they participate in your research project, and with regard to any and all data that may be collected from the subjects. The Board's approval includes continued monitoring of your research by the Board to assure that the subjects are being adequately and properly protected from such risks. If at any time a subject becomes injured or complains of injury, you must notify Serena Stanford, Ph.D., immediately. Injury includes but is not limited to bodily harm, psychological trauma and release of potentially damaging personal information.

Please also be advised that all subjects need to be fully informed and aware that their participation in your research project is voluntary, and that he or she may withdraw from the project at any time. Further, a subject's participation, refusal to participate, or withdrawal will not affect any services the subject is receiving or will receive at the institution in which the research is being conducted.

If you have any questions, please contact me at (408) 924-2480.

The California State University:  
Chancellor's Office  
Bakersfield, Chico, Dominguez Hills  
Fresno, Fullerton, Hayward, Humboldt  
Long Beach, Los Angeles, Maritime Academy  
Monterey Bay, Northridge, Pomona  
Sacramento, San Bernardino, San Diego  
San Francisco, San José, San Luis Obispo  
San Marcos, Sonoma, Stanislaus